

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 11:40:16 ; Search time 62 Seconds  
(without alignments)  
184.832 Million cell updates/sec

Title: PCT-US02-27855-1

Perfect score: 499  
Sequence: 1 KTCYEGNGHFYRGKASIDTM.....YVGVGLKPLVQECNVHDCAD 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	494	99.0	88	23	AAE16542
2	494	99.0	96	23	AAE16550
3	494	99.0	135	23	AAE16545
4	494	99.0	143	23	AAE16549
5	494	99.0	337	22	AAE16542
6	494	99.0	337	23	ABP41795
7	494	99.0	365	16	AAE68854
8	494	99.0	378	11	AAW13635
9	494	99.0	386	16	AAE6286
10	494	99.0	389	11	AAW13636

11	494	99.0	390	16	AAE6245	Bifunctional uroki
12	494	99.0	390	16	AAE6247	Bifunctional uroki
13	494	99.0	392	16	AAE6254	Bifunctional uroki
14	494	99.0	392	16	AAE6255	Bifunctional uroki
15	494	99.0	392	16	AAE6256	Bifunctional uroki
16	494	99.0	392	16	AAE6257	Bifunctional uroki
17	494	99.0	392	16	AAE6258	Bifunctional uroki
18	494	99.0	392	16	AAE6259	Bifunctional uroki
19	494	99.0	392	16	AAE6260	Bifunctional uroki
20	494	99.0	392	16	AAE6261	Bifunctional uroki
21	494	99.0	392	16	AAE6263	Bifunctional uroki
22	494	99.0	392	16	AAE6264	Bifunctional uroki
23	494	99.0	393	16	AAE6244	Bifunctional uroki
24	494	99.0	393	16	AAE6249	Bifunctional uroki
25	494	99.0	393	16	AAE6250	Bifunctional uroki
26	494	99.0	393	16	AAE6251	Bifunctional uroki
27	494	99.0	393	16	AAE6252	Bifunctional uroki
28	494	99.0	393	16	AAE6253	Bifunctional uroki
29	494	99.0	393	17	AAE99885	M36: fibrinolytic
30	494	99.0	393	17	AAE99596	Chimeric protein M
31	494	99.0	393	17	AAE9597	Pro-urokinase deri
32	494	99.0	395	15	AAE47902	Bifunctional uroki
33	494	99.0	395	16	AAE6252	Bifunctional uroki
34	494	99.0	395	16	AAE6265	Bifunctional uroki
35	494	99.0	396	16	AAE6246	Bifunctional uroki
36	494	99.0	397	16	AAE6248	Bifunctional uroki
37	494	99.0	401	11	AAW13637	Human prourokinase
38	494	99.0	403	23	AAE16547	Human urokinase-ty
39	494	99.0	410	13	AAE23794	Prourokinase mutan
40	494	99.0	410	13	AAE23795	Prourokinase doubl
41	494	99.0	411	6	AAE50871	Sequence encoded b
42	494	99.0	411	9	AAE80428	Sequence of mutant
43	494	99.0	411	11	AAW13634	Human native prour
44	494	99.0	411	11	AAE04185	Plasminogen activa
45	494	99.0	411	11	AAE05117	UK-53 as encoded b

#### ALIGNMENTS

RESULT 1	AAE16542	AAE16542 standard; Protein: 88 AA.
ID	AAE16542	
XX	AAE16542;	
AC		
XX		
DT	09-APR-2002 (first entry)	
XX		
DE	Human urokinase-type plasminogen activator (uPA) kringle.	
XX		
KW	Human: urokinase-type plasminogen activator; uPA; therapy: hypertension;	
KW	stroke: hypotension; atherosclerosis; heart attack; thrombotic disorder;	
KW	microvascular occlusion; angiodysplasia; pulmonary fibrosis; asthma;	
KW	tumor cell metastasis; glaucoma; diabetic retinopathy; wound healing;	
KW	clotting disorder; uterine contraction disorder; respiratory disease;	
KW	male impotence; adult respiratory distress syndrome.	
OS	Homo sapiens.	
XX		
PN	WO200197752-A2.	
XX		
PD	27-DEC-2001.	
XX		
PF	13-JUN-2001; 2001WO-US18976.	
XX		
PR	20-JUN-2000; 2000US-212874P.	
XX		
PA	(UYPE-) UNTV PENNSYLVANIA.	
XX		
PI	Cines DB, Hlgazi AA;	
XX		
DR	WPI: 2002-122240/16.	
	N-PSDB: AAD27075.	

XX Composition for modulating muscle cell and tissue contractility for  
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
 PT comprising domains from urokinase-type plasminogen activator -  
 PS Claim 1, Fig 1A; 117pp; English.

XX The invention relates to a composition comprising one or more domains of  
 CC urokinase-type plasminogen activator (uPA). The composition is used to  
 CC modulate the contractility and angiogenic activity of a mammalian muscle,  
 CC endothelial cell or tissue. The composition is used for treating stroke,  
 CC hypertension, atherosclerosis, heart attack, microvascular  
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
 CC disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell  
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
 CC diabetic retinopathy, wound healing, clotting disorder, uterine  
 CC contraction disorder, male impotence, respiratory disease or condition  
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
 CC hypertension, microvascular thrombotic occlusion, and a disorder  
 CC associated with chronic intrapulmonary fibrin formation. The present  
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle.  
 XX

Sequence 88 AA:

Query Match 99.0%; Score 494; DB 23; Length 88;  
 Best Local Similarity 98.8%; Pred. No. 1.8e-41;  
 Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLAQOTYHAHRSNALQGLGKHNYCRNPN 60  
 DB 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLAQOTYHAHRSDALQGLGKHNYCRNPN 60  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86  
 DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 2  
 AAE16550  
 ID AAE16550 standard; Protein; 96 AA.

AC AAE16550;

DT 09-APR-2002 (first entry)

XX Human uPA kringle and connecting peptide.

XX Human; urokinase-type plasminogen activator; uPA; therapy: hypertension;  
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
 KW microvascular occlusion; angiotensin disorder; pulmonary fibrosis; asthma;  
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
 KW clotting disorder; uterine contraction disorder; respiratory disease;  
 KW adult respiratory distress syndrome; male impotence.

XX Homo sapiens.

PN WO200197752-A2.

PD 27-DEC-2001.

PF 13-JUN-2001; 2001WO-US18976.

PR 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

PI Cines DB, Higazi AA;

DR WPI: 2002-122240/16.

DR N-PSDB: AAD27083.

PT Composition for modulating muscle cell and tissue contractility for  
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,

PT comprising domains from urokinase-type plasminogen activator -  
 PS Claim 25; Fig 11; 117pp; English.

XX The invention relates to a composition comprising one or more domains of  
 CC urokinase-type plasminogen activator (uPA). The composition is used to  
 CC modulate the contractility and angiogenic activity of a mammalian muscle,  
 CC endothelial cell or tissue. The composition is used for treating stroke,  
 CC hypertension, atherosclerosis, heart attack, microvascular  
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
 CC disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell  
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
 CC diabetic retinopathy, wound healing, clotting disorder, uterine  
 CC contraction disorder, male impotence, respiratory disease or condition  
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
 CC hypertension, microvascular thrombotic occlusion, and a disorder  
 CC associated with chronic intrapulmonary fibrin formation. The present  
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle  
 CC and connecting peptide.  
 XX

Sequence 96 AA:

Query Match 99.0%; Score 494; DB 23; Length 96;  
 Best Local Similarity 98.8%; Pred. No. 1.9e-41;  
 Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLAQOTYHAHRSNALQGLGKHNYCRNPN 60  
 DB 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLAQOTYHAHRSDALQGLGKHNYCRNPN 60  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86  
 DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 3  
 AAE16545  
 ID AAE16545 standard; Protein; 135 AA.

AC AAE16545;

DT 09-APR-2002 (first entry)

XX Human urokinase-type plasminogen activator amino terminal fragment (ATF).

XX Human; urokinase-type plasminogen activator; uPA; therapy: hypertension;  
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
 KW microvascular occlusion; angiotensin disorder; pulmonary fibrosis; asthma;  
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
 KW clotting disorder; uterine contraction disorder; respiratory disease;  
 KW adult respiratory distress syndrome; amino terminal fragment; ATF;  
 KW male impotence.

XX Homo sapiens.

PN WO200197752-A2.

PD 27-DEC-2001.

PF 13-JUN-2001; 2001WO-US18976.

PR 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

PI Cines DB, Higazi AA;

DR WPI: 2002-122240/16.

DR N-PSDB: AAD27078.

PT Composition for modulating muscle cell and tissue contractility for  
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
 PT comprising domains from urokinase-type plasminogen activator -

XX Claim 11; Fig 1D; 117pp; English.

XX The invention relates to a composition comprising one or more domains of  
XX urokinase-type plasminogen activator (uPA). The composition is used to  
XX modulate the contractility and angiogenic activity of a mammalian muscle,  
XX endothelial cell or tissue. The composition is used for treating stroke,  
XX hypertension, atherosclerosis, heart attack, microvascular  
XX occlusions, thrombotic microangiopathies, surgically induced thrombotic  
XX disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell  
XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
XX diabetic retinopathy, wound healing, clotting disorder, uterine  
XX contraction disorder, male impotence, respiratory disease or condition  
XX such as asthma, adult respiratory distress syndrome, primary pulmonary  
XX hypertension, microvascular thrombotic occlusion, and a disorder  
XX associated with chronic intrapulmonary fibrin formation. The present  
XX sequence is human urokinase-type plasminogen activator (uPA) amino  
XX terminal fragment (ATF).

XX Sequence 135 AA;

XX Query Match

Best Local Similarity 99.0%; Score 494; DB 23; Length 135;

XX Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTPMGRPCLPMSATVLOQTYHAHRSNALQLGLGKHNCRPN 60

DB 48 KTCYEGNGHFYRGKASTDTPMGRPCLPMSATVLOQTYHAHRSNALQLGLGKHNCRPN 107

QY 61 RRRPWCYVOYGLKPLVQECNVHDCAD 86

DB 108 RRRPWCYVOYGLKPLVQECNVHDCAD 133

RESULT 4

AAE16549 ID AAE16549 standard; Protein; 143 AA.

XX AAE16549;

XX 09-APR-2002 (first entry)

XX Human uPA amino terminal fragment (ATF) and connecting peptide.

XX Human: urokinase-type plasminogen activator; uPA; therapy; hypertension;  
XX stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
XX microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
XX tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
XX clotting disorder; uterine contraction disorder; respiratory disease;  
XX adult respiratory distress syndrome; amino terminal fragment; ATF;  
XX male impotence.

XX Homo sapiens.

XX WO200197752-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

XX N-PSDB; AAD27082.

XX Composition for modulating muscle cell and tissue contractility for  
XX treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
XX comprising domains from urokinase-type plasminogen activator -

PS Claim 24; Fig 1H; 117pp; English.

XX The invention relates to a composition comprising one or more domains of  
XX urokinase-type plasminogen activator (uPA). The composition is used to  
XX modulate the contractility and angiogenic activity of a mammalian muscle,  
XX endothelial cell or tissue. The composition is used for treating stroke,  
XX hypertension, atherosclerosis, heart attack, microvascular  
XX occlusions, thrombotic microangiopathies, surgically induced thrombotic  
XX disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell  
XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
XX diabetic retinopathy, wound healing, clotting disorder, uterine  
XX contraction disorder, male impotence, respiratory disease or condition  
XX such as asthma, adult respiratory distress syndrome, primary pulmonary  
XX hypertension, microvascular thrombotic occlusion, and a disorder  
XX associated with chronic intrapulmonary fibrin formation. The present  
XX sequence is human urokinase-type plasminogen activator (uPA) amino  
XX terminal fragment (ATF) and connecting peptide.

XX Sequence 143 AA;

XX Query Match

Best Local Similarity 99.0%; Score 494; DB 23; Length 143;

XX Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTPMGRPCLPMSATVLOQTYHAHRSNALQLGLGKHNCRPN 60

DB 48 KTCYEGNGHFYRGKASTDTPMGRPCLPMSATVLOQTYHAHRSNALQLGLGKHNCRPN 107

QY 61 RRRPWCYVOYGLKPLVQECNVHDCAD 86

DB 108 RRRPWCYVOYGLKPLVQECNVHDCAD 133

RESULT 5

AAG75492 ID AAG75492 standard; Protein; 337 AA.

XX AAG75492;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:6256.

XX Human: colon cancer; colon cancer antigen; diagnosis; detection;  
XX colorectal carcinoma; chromosome 10.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAH34897.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 7707-7708; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
XX cancer-associated nucleic acid molecules (N) and proteins (P), where  
XX the proteins are collectively known as colon cancer antigens. The colon  
XX cancer antigens have cytosolic activity and can be used in gene

CC therapy and vaccine production. N and P may be used in the prevention,  
 CC diagnosis and treatment of diseases associated with inappropriate P  
 CC expression. For example, N and P may be used to treat disorders  
 CC associated with decreased expression by rectifying mutations or deletions  
 CC in a patient's genome that affect the activity of P by expressing P.  
 CC Additionally, N may be used to produce the colon cancer-associated Ps,  
 CC by inserting the nucleic acids into a host cell and culturing the cell  
 CC to express the proteins. N and P can be used in the prevention, diagnosis  
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
 CC and AAB77789 represent sequences used in the exemplification of the  
 CC present invention.  
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
 CC missing at time of publication, meaning no sequences are present for  
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 337 AA:

Query Match 99.0%: Score 494; DB 23; Length 337;  
 Best Local Similarity 98.8%; Pred. No. 6, 5e-41;  
 Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASDTDMGRPCLPWNSATVYLQOTYHAHRSNALQGLGKHNHCNPND 60  
 DB 74 KTCYEGNGHFYRGKASDTDMGRPCLPWNSATVYLQOTYHAHRSNALQGLGKHNHCNPND 133  
 OY 61 RRRPWCYVQGLKPLVQECVHDCAD 86  
 DB 134 RRRPWCYVQGLKPLVQECVHDCAD 159

RESULT 6  
 AAB41795  
 ID ABA41795 standard; Protein; 337 AA.

XX ABA41795;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HVCB879, SEQ ID NO:2927.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive; chromosome 10q24.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI: 2002-147878/19.

XX N-PSDB: ABQ54872.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
 PT ovarian cancer), immune disorders, cardiovascular disorders and  
 PT neurological diseases -

XX Claim 11; SEQ ID NO 2927; 2922pp; English.

PS The invention relates to 2175 novel human ovarian antigens (AAB41054-  
 CC ABA43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing  
 CC treating, prognosis or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp://pub/published\_pct\_sequences.

XX Sequence 337 AA:

Query Match 99.0%: Score 494; DB 23; Length 337;  
 Best Local Similarity 98.8%; Pred. No. 6, 5e-41;  
 Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASDTDMGRPCLPWNSATVYLQOTYHAHRSNALQGLGKHNHCNPND 60  
 DB 74 KTCYEGNGHFYRGKASDTDMGRPCLPWNSATVYLQOTYHAHRSNALQGLGKHNHCNPND 133  
 OY 61 RRRPWCYVQGLKPLVQECVHDCAD 86  
 DB 134 RRRPWCYVQGLKPLVQECVHDCAD 159

RESULT 7  
 AAR68854  
 ID AAR68854 standard; protein; 365 AA.

XX AAR68854;

XX 22-NOV-1995 (first entry)

XX Delta 1-46 urokinase.

XX Human; des-epidermal growth factor homologous plasminogen activator;  
 KW uPA; liver membrane; reduced affinity; BGF homologous; thrombolysis;  
 KW thrombolytic; increased half-life; urokinase.

XX Homo sapiens (engineered).

XX Key Location/Qualifiers

XX MISC-difference 1 /note="amino acids 1-46 of wild-type urokinase  
 XX have been deleted"

XX US5376547-A.

XX 27-DEC-1994.

PF 30-JAN-1987; 87US-0008795.  
 XX 30-JAN-1987; 87US-0008795.  
 PR 29-JAN-1988; 88US-0150267.  
 XX  
 PA (AMHP) AMERICAN HOME PROD CORP.  
 XX  
 PI Hung PP, Kalyan NK, Lee SL;  
 XX WPI: 1995-043464/06.  
 DR  
 XX New modified plasminogen activator cpds. - having regions removed  
 PT to reduce affinity for liver membranes and increase circulation  
 PR half-life.  
 XX  
 PS Claim 1; 26pp; English.  
 XX  
 CC Amino acid residues 1-46 contain the EGF region of human urokinase.  
 CC Deletion of this region results in a plasminogen activator with  
 CC reduced affinity for liver cell membranes; the mutant protein is  
 CC not cleared from the circulation as rapidly as its wild-type tpa.  
 CC The specification only gives the sequence around the deletion and  
 CC not the full-length sequence of "delta 1-46 urokinase"; the  
 CC sequence in AR68854 has been obtained by amending a previously  
 CC disclosed wild-type human urokinase sequence (from WO9501427)  
 CC according to the description given in Example 3.  
 CC  
 XX  
 SQ Sequence 365 AA;  
 Query Match 99.0%; Score 494; DB 16; Length 365;  
 Best Local Similarity 98.8%; Pred. No. 7e-41;  
 Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOOTYHAHNSNALQLGLGKHNCRPN 60  
 DB 2 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOOTYHAHNSDALQLGLGKHNCRPN 61  
 OY 61 RRRPWCYOVGLKPLVQECMVHDCAD 86  
 DB 62 RRRPWCYOVGLKPLVQECMVHDCAD 87  
 RESULT 8  
 AAW13635  
 ID AAW13635 standard; Protein; 378 AA.  
 XX  
 AC AAW13635;  
 XX  
 DT 04-JUN-1997 (first entry)  
 XX  
 DE Human prourokinase variant lacking entire EGF domain.  
 XX  
 KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;  
 KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..9  
 FT /note= "residues 1-9 of native hPUK"  
 FT Region 10..378  
 FT /note= "residues 43-411 of native hPUK"  
 FT MISC-difference 144  
 FT /note= "corresponds to TAC codon"  
 XX  
 PN EP398361-A.  
 XX  
 PD 22-NOV-1990.  
 XX  
 PF 18-MAY-1990; 90EP-0109472.  
 XX  
 PR 22-FEB-1990; 90JP-0042020.

PR 18-MAY-1989; 89JP-0126433.  
 PR 03-JUL-1986; 86JP-0156936.  
 PR 18-FEB-1987; 87JP-0036495.  
 PR 18-MAY-1989; 89JP-0126434.  
 XX  
 PA (GREG) GREEN CROSS CORP.  
 XX  
 PI Altmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;  
 PI Morita M, Tanabe T;  
 XX WPI: 1990-350146/47.  
 DR N-PSDB: AAT61672.  
 XX  
 PT Human pro-urokinase variants - deficient in loop regions of  
 PT epidermal growth factor, showing long blood half-life, as  
 PT fibrinolytic agent  
 XX  
 PS Claim 1; Page -: 22pp; English.  
 XX  
 CC New variants of human prourokinase (hPUK) comprise a hPUK deficient  
 CC in (i) at least part of the first loop region of the epidermal growth  
 CC factor (EGF) domain; (ii) at least part of the first loop and at  
 CC least part of the second loop; or (iii) at least part of the third  
 CC loop. The hPUK variants show an increased blood half-life comparable  
 CC to that of the whole EGF domain-deficient hPUK variant and urokinase  
 CC while retaining the same properties as those of hPUK. They have  
 CC potent thrombolytic activity and very little tendency to cause  
 CC spontaneous bleeding. The present sequence represents a specific variant  
 CC of hPUK which lacks the entire EGF domain; the sequence does not  
 CC appear in the specification and has been created using the  
 CC wild-type hPUK sequence and the junction sequence after  
 CC deletion, both of which are given (in Fig 1 and in Fig 2(3),  
 CC respectively).  
 XX  
 SQ Sequence 378 AA;  
 Query Match 99.0%; Score 494; DB 11; Length 378;  
 Best Local Similarity 98.8%; Pred. No. 7.2e-41;  
 Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOOTYHAHNSNALQLGLGKHNCRPN 60  
 DB 15 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOOTYHAHNSDALQLGLGKHNCRPN 74  
 OY 61 RRRPWCYOVGLKPLVQECMVHDCAD 86  
 DB 75 RRRPWCYOVGLKPLVQECMVHDCAD 100  
 RESULT 9  
 AAR6266  
 ID AAR6266 standard; Protein; 386 AA.  
 XX  
 AC AAR6266;  
 XX  
 DT 17-AUG-1995 (first entry)  
 XX  
 DE Bifunctional urokinase variant M33.  
 XX  
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"  
 FT Disulfide-bond 4..85  
 FT Disulfide-bond 25..67  
 FT Disulfide-bond 56..80  
 FT Disulfide-bond 102..223  
 FT Disulfide-bond 143..159

FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Region 365..366  
FT /label= X1  
FT /note= "peptide bond"  
FT Region 366..386  
FT /label= Y1  
PN DE4323754-C.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-4323754.  
XX 15-JUL-1993; 93DE-4323754.  
XX (CHEF ) GRUENTHAL GMBH.  
PA Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
PI Wment S;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with  
PT improved fibrinolytic and thrombin inhibiting activities, for  
PT treating cardiac and cerebral infarct, pulmonary embolism, etc  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX CC Bifunctional urokinase derivatives corresponding to the formula  
CC M4-XI-1 are claimed (see features table). Sequences AAR6244-R66266  
CC are specific examples of such derivs. which have both improved  
CC fibrinolytic and thrombin-inhibiting activities, compared to known  
CC plasminogen activators or thrombin inhibitors. The proteins are  
CC useful as thrombolytic agents, e.g. for treatment of arterial  
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction  
CC and pulmonary embolism.  
XX CC  
XX SQ Sequence 386 AA;  
XX  
XX Query Match 99.0%; Score 494; DB 16; Length 386;  
XX Best Local Similarity 98.8%; Pred. No. 7.4e-41;  
XX Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFRKASTDTMGRCPLPWNASATVLDQTYHAHRSNALQGLGKHNCRPN 60  
DB 2 KTCYEGNGHFRKASTDTMGRCPLPWNASATVLDQTYHAHRSNALQGLGKHNCRPN 61  
QY 61 RRRPWCYVQGLKPLVDECMVHDCAD 86  
DB 62 RRRPWCYVQGLKPLVDECMVHDCAD 87  
RESULT 10  
AAW13636  
ID AAW13636 standard; Protein; 389 AA.  
XX  
XX AAW13636;  
XX  
XX 04-JUN-1997 (first entry)  
XX Human prourokinase variant lacking EGF domain loops 1 and 2.  
XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;  
KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Region 1..10

FT Region /note= "residues 1-10 of native hPUK"  
FT 11..389  
FT /note= "residues 33-411 of native hPUK"  
FT Misc-difference 155  
FT /note= "corresponds to TAC codon"  
XX EP398361-A.  
XX  
XX 22-NOV-1990.  
XX  
XX 18-MAY-1990; 90EP-0109472.  
XX  
XX 22-FEB-1990; 90JP-0042020.  
XX 18-MAY-1989; 89JP-0126433.  
XX 03-JUL-1986; 86JP-0156936.  
XX 18-FEB-1987; 87JP-0036495.  
XX 18-MAY-1989; 89JP-0126434.  
XX  
XX (GREC ) GREEN CROSS CORP.  
XX  
XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;  
PI Morita M, Tanabe T;  
XX WPI; 1990-350146/47.  
XX N-PSDB; AAT61673.  
XX  
XX Human pro-urokinase variants - deficient in loop regions of  
PT epidermal growth factor, showing long blood half-life, as  
PT fibrinolytic agent  
XX  
XX Claim 6; Page -; 22pp; English.  
XX  
XX New variants of human prourokinase (hPUK) comprise a hPUK deficient  
CC in (1) at least part of the first loop region of the epidermal growth  
CC factor (EGF) domain; (11) at least part of the first loop and at  
CC least part of the second loop; or (11) at least part of the third  
CC loop. The hPUK variants show an increased blood half-life comparable  
CC to that of the whole EGF domain-deficient hPUK variant and urokinase  
CC while retaining the same properties as those of hPUK. They have  
CC potent thrombolytic activity and very little tendency to cause  
CC spontaneous bleeding. The present sequence represents a specific  
CC variant of hPUK which lacks loops 1 and 2 of the EGF domain; the  
CC sequence does not appear in the specification and has been created  
CC using the wild-type hPUK sequence and the junction sequence after  
CC deletion, both of which are given (in Fig 1 and on page 8,  
CC respectively).  
XX CC  
XX SQ Sequence 389 AA;  
XX  
XX Query Match 99.0%; Score 494; DB 11; Length 389;  
XX Best Local Similarity 98.8%; Pred. No. 7.5e-41;  
XX Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFRKASTDTMGRCPLPWNASATVLDQTYHAHRSNALQGLGKHNCRPN 60  
DB 26 KTCYEGNGHFRKASTDTMGRCPLPWNASATVLDQTYHAHRSNALQGLGKHNCRPN 85  
QY 61 RRRPWCYVQGLKPLVDECMVHDCAD 86  
DB 86 RRRPWCYVQGLKPLVDECMVHDCAD 111  
RESULT 11  
AAR66245  
ID AAR66245 standard; Protein; 390 AA.  
XX  
XX AAR66245;  
XX  
XX 17-AUG-1995 (first entry)  
XX Bifunctional urokinase variant M12.  
XX  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

DT	17-AUG-1995	(first entry)
XX		
DE	Bifunctional urokinase variant M14.	
XX		
KW	fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;	
XX	urokinase; variant; mutein.	
XX		
OS	Synthetic.	
XX		
Key	Location/Qualifiers	
Region	1..365	
FT	/Label= M4	
FT	/note= "unglycosylated prourokinase( Ser47-Leu411) "	
FT	Disulfide-bond	4..85
FT	Disulfide-bond	25..67
FT	Disulfide-bond	56..80
FT	Disulfide-bond	102..233
FT	Disulfide-bond	143..159
FT	Disulfide-bond	151..222
FT	Disulfide-bond	247..316
FT	Disulfide-bond	279..295
FT	Disulfide-bond	306..334
FT	Disulfide-bond	366..371
Region	/Label= X1	
FT	/Label= X1	
FT	372..390	
FT	/Label= Y1	
XX		
XX	DE4323754-C.	
PD	01-DEC-1994.	
XX		
XX	15-JUL-1993; 93DE-4323754.	
XX		
XX	15-JUL-1993; 93DE-4323754.	
PR		
XX		
PA	(CHEF ) GRUENTHAL GMBH.	
XX		
PI	Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;	
PI	Wendt S;	
XX		
DR	WPI; 1995-015191/03.	
XX		
PT	New bifunctional urokinase derivs and related plasmids - with	
PT	improved fibrinolytic and thrombin inhibiting activities, for	
PT	treating cardiac and cerebral infarct, pulmonary embolism, etc	
XX		
PS	Example 1; Page 10 and Fig 1; 34pp; German.	
XX		
CC	Bifunctional urokinase derivatives corresponding to the formula	
CC	M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266	
CC	are specific examples of such derivs. which have both improved	
CC	fibrinolytic and thrombin-inhibiting activities, compared to known	
CC	plasmaogen activators or thrombin inhibitors. The proteins are	
CC	useful as thrombolytic agents, e.g. for treatment of arterial	
CC	occlusions, deep vein thrombosis, cardiac and cerebral infarction	
CC	and pulmonary embolism.	
XX		
XX		
Sequence	390 AA:	
Query Match	99.0%; Score 494; DB 16; Length 390;	
Best Local Similarity	98.8%; Pred. No. 7.5e-41;	
Matches	85; Conservative 1; Mismatches 0; Indels 0; Gaps 0	
DB	1 KTCFEGNGHYRKRASDPTMGRCPLPMNSATYLAQTYHAHRSNMLQGLGKHNYCRPN 60	
	2 KTCFEGNGHYRKRASDPTMGRCPLPMNSATYLAQTYHAHRSNMLQGLGKHNYCRPN 61	
QY	61 RRRPWCYVOYGLKPLVQECVHDCAD 86	
	62 RRRPWCYVOYGLKPLVQECVHDCAD 87	
DB		
RESULT	13	

```

AAR66254
ID AAR66254 standard; protein; 392 AA.
XX
AC AAR66254;
XX
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M21.
XX
KM fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key
FH Location/Qualifiers
FH 1..365
FH Region
FH Disulfide-bond
FT 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..295
FT Disulfide-bond 279..316
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region
FT 372..392
FT /label= X1
FT /label= Y1
XX
XX DE4323754-C.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wmendt S;
XX WPI; 1995-015191/03.
XX
XX DR
XX
XX PT New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX
XX SQ Sequence 392 AA;
XX
XX Query Match 99.0%; Score 494; DB 16; Length 392;
XX Best Local Similarity 98.8%; Pred. No. 7.5e-41;
XX Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 62 RRRPMCYOVGGLKPLVQECMVHDCAD 87
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
RESULT 14
AAR66255
ID AAR66255 standard; protein; 392 AA.
XX
AC AAR66255;
XX
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M22.
XX
KM fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key
FH Location/Qualifiers
FH 1..365
FH Region
FH Disulfide-bond
FT 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region
FT 372..392
FT /label= X1
FT /label= Y1
XX
XX DE4323754-C.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wmendt S;
XX WPI; 1995-015191/03.
XX
XX DR
XX
XX PT New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX
XX SQ Sequence 392 AA;
XX
XX Query Match 99.0%; Score 494; DB 16; Length 392;
XX Best Local Similarity 98.8%; Pred. No. 7.5e-41;
XX Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```



QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSNALQGLGKHNCRPN 60  
 DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSNALQGLGKHNCRPN 61  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86  
 DB 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87

## RESULT 15

AAR66256  
 ID AAR66256 standard; protein; 392 AA.  
 AC AAR66256;

DT 17-AUG-1995 (first entry)  
 XX

DE Bifunctional urokinase variant M23.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutin.  
 XX

OS Synthetic.

Key Location/Qualifiers  
 FH 1..365  
 FT /label= M4

FT /note= "unglycosylated prourokinase(Ser47-Leu411)"  
 FT 4..85

FT Disulfide-bond 25..67

FT Disulfide-bond 56..80

FT Disulfide-bond 102..233

FT Disulfide-bond 143..159

FT Disulfide-bond 151..222

FT Disulfide-bond 247..316

FT Disulfide-bond 279..295

FT Disulfide-bond 306..334

FT Disulfide-bond 366..371

FT /label= X1

FT Region 372..392

FT /label= Y1

PN DE4323754-C.

PD 01-DEC-1994.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 99.0%; Score 494; DB 16; Length 392;  
 Best Local Similarity 98.8%; Pred. No. 7.5e-41;  
 Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSNALQGLGKHNCRPN 60  
 DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSNALQGLGKHNCRPN 61

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86  
 DB 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87

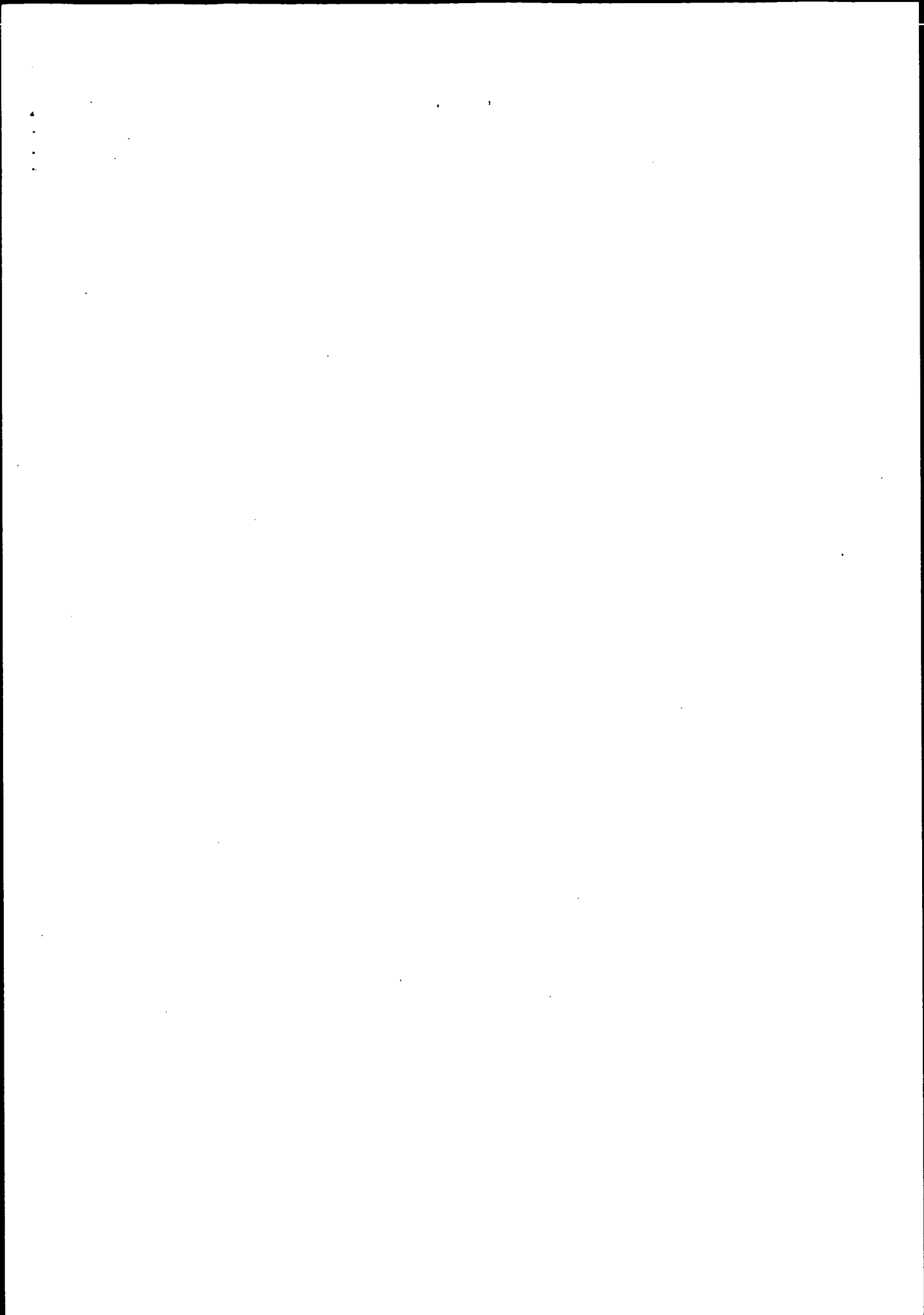
Search completed: November 11, 2002, 12:59:49  
 Job time : 62 secs

New bifunctional urokinase derivs and related plasmids - with improved fibrinolytic and thrombin inhibiting activities, for treating cardiac and cerebral infarct, pulmonary embolism, etc

Example 1; Page 10 and Fig 1; 34pp; German.

Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266 are specific examples of such derivs. which have both improved fibrinolytic and thrombin-inhibiting activities, compared to known plasminogen activators or thrombin inhibitors. The proteins are useful as thrombolytic agents, e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac and cerebral infarction and pulmonary embolism.

Sequence 392 AA;



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 12:58:41 ; Search time 14 Seconds  
(without alignments)  
88.571 Million cell updates/sec

Title: PCT-US02-27855-1

Perfect score: 499

Sequence: 1 KTCYGNHGFYRGKASTDTM.....YVQVGLKPLVQECMWHDCAD 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 92612 seqs, 14418503 residues

Total number of hits satisfying chosen parameters: 92612

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCUTS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	99.0	88	US-09-880-503-1	Sequence 1, Appli
2	494	99.0	96	US-09-880-503-9	Sequence 9, Appli
3	494	99.0	135	US-09-880-503-4	Sequence 4, Appli
4	494	99.0	138	US-09-880-503-8	Sequence 12, Appli
5	494	99.0	143	US-09-880-503-6	Sequence 8, Appli
6	494	99.0	403	US-09-880-503-3	Sequence 6, Appli
7	494	99.0	411	US-09-880-503-3	Sequence 3, Appli
8	494	99.0	431	US-09-264-4688-1	Sequence 1, Appli
9	216	43.3	372	US-09-084-491A-3	Sequence 3, Appli
10	216	43.3	562	US-09-969-271-7	Sequence 7, Appli
11	215	43.1	562	US-09-974-298-145	Sequence 145, App
12	188.5	37.8	560	US-09-912-559-3	Sequence 3, Appli
13	188.5	37.8	560	US-09-912-559-4	Sequence 4, Appli
14	151	30.3	368	US-09-761-120-42	Sequence 42, Appli
15	149	29.9	378	US-09-873-676-1	Sequence 1, Appli
16	149	29.9	458	US-09-946-893-4	Sequence 4, Appli
17	149	29.9	569	US-09-946-893-5	Sequence 5, Appli
18	149	29.9	571	US-09-946-893-8	Sequence 8, Appli
19	149	29.9	576	US-09-946-893-6	Sequence 6, Appli

20	149	29.9	791	US-09-967-386-1	Sequence 1, Appli
21	149	29.9	810	US-09-946-893-2	Sequence 2, Appli
22	148.5	29.8	78	US-09-753-064-5	Sequence 5, Appli
23	148.5	29.8	78	US-09-761-120-23	Sequence 23, Appli
24	148.5	29.8	352	US-09-761-120-40	Sequence 40, Appli
25	147	29.5	937	US-09-974-298-129	Sequence 129, App
26	137	27.5	378	US-09-761-120-41	Sequence 41, Appli
27	137	27.5	459	US-09-761-120-46	Sequence 46, Appli
28	137	27.5	812	US-09-788-142-1	Sequence 1, Appli
29	137	27.5	812	US-09-761-120-1	Sequence 1, Appli
30	137	27.5	812	US-09-873-676-81	Sequence 81, Appli
31	136.5	27.4	339	US-09-788-142-3	Sequence 3, Appli
32	136.5	27.4	339	US-09-788-142-6	Sequence 6, Appli
33	136.5	27.4	339	US-09-761-120-3	Sequence 3, Appli
34	136.5	27.4	339	US-09-761-120-6	Sequence 6, Appli
35	135.5	27.2	78	US-09-761-120-22	Sequence 22, Appli
36	135.5	27.2	352	US-09-761-120-39	Sequence 39, Appli
37	134.5	27.0	339	US-09-788-142-5	Sequence 5, Appli
38	134.5	27.0	339	US-09-761-120-5	Sequence 5, Appli
39	133.5	26.8	339	US-09-788-142-4	Sequence 4, Appli
40	133.5	26.8	339	US-09-761-120-4	Sequence 4, Appli
41	132	26.5	72	US-09-848-288-6	Sequence 6, Appli
42	132	26.5	91	US-09-748-468-1	Sequence 1, Appli
43	132	26.5	263	US-09-084-491A-2	Sequence 2, Appli
44	132	26.5	339	US-09-788-142-2	Sequence 2, Appli
45	132	26.5	339	US-09-761-120-2	Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-09-880-503-1

Sequence 1, Application US/09880503  
Patent No. US20020131964A1

GENERAL INFORMATION:

APPLICANT: CINES, Douglas B

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

TITLE OF INVENTION: TISSUE CONTRACTILITY

FILE REFERENCE: 9596-331

CURRENT APPLICATION NUMBER: US/09/880, 503

PRIOR FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/212, 847

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 88

TYPE: PRT

ORGANISM: Homo sapiens

US-09-880-503-1

Query Match 99.0%; Score 494; DB 10; Length 88;  
Best Local Similarity 98.8%; Pred. No. 5.4e-51;

Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYGNHGFYRGKASTDTMGRPCLPMSATVLTQTYAHHSNMLQIGCKHNYCRPDN 60  
|||||  
DB 1 KTCYGNHGFYRGKASTDTMGRPCLPMSATVLTQTYAHHSNMLQIGCKHNYCRPDN 60

QY 61 RRRPWCYQVGLKPLVQECMWHDCAD 86  
|||||  
DB 61 RRRPWCYQVGLKPLVQECMWHDCAD 86

RESULT 2  
US-09-880-503-9

Sequence 9, Application US/09880503  
Patent No. US20020131964A1

GENERAL INFORMATION:

APPLICANT: CINES, Douglas B

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

TITLE OF INVENTION: TISSUE CONTRACTILITY

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
;; FILE OF INVENTION: TISSUE CONTRACTABILITY  
;; FILE REFERENCE: 9596-331  
;; CURRENT APPLICATION NUMBER: US/09/880,503  
;; CURRENT FILING DATE: 2001-06-13  
;; PRIOR APPLICATION NUMBER: US 60/212,847  
;; PRIOR FILING DATE: 2000-06-20  
;; NUMBER OF SEQ ID NOS: 18  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO: 9  
;; LENGTH: 96  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-880-503-9

Query Match 99.0%; Score 494; DB 10; Length 96;  
Best Local Similarity 98.8%; Pred. No. 5.9e-51;  
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNCRPN 60  
DB 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNCRPN 60  
QY 61 RRRPWCYVQVGLKPLVOECMVHDCAD 86  
DB 61 RRRPWCYVQVGLKPLVOECMVHDCAD 86

RESULT 3  
US-09-880-503-4  
;; Sequence 4 Application US/09880503  
;; Patent No. US20020131964A1  
;; GENERAL INFORMATION:  
;; APPLICANT: CINES, Douglas B  
;; APPLICANT: HIGAZI, Abd Al-Rouf  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
;; FILE REFERENCE: 9596-331  
;; CURRENT APPLICATION NUMBER: US/09/880,503  
;; CURRENT FILING DATE: 2001-06-13  
;; PRIOR APPLICATION NUMBER: US 60/212,847  
;; PRIOR FILING DATE: 2000-06-20  
;; NUMBER OF SEQ ID NOS: 18  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO: 4  
;; LENGTH: 135  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-880-503-4

Query Match 99.0%; Score 494; DB 10; Length 135;  
Best Local Similarity 98.8%; Pred. No. 8.4e-51;  
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNCRPN 60  
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNCRPN 107  
QY 61 RRRPWCYVQVGLKPLVOECMVHDCAD 86  
DB 108 RRRPWCYVQVGLKPLVOECMVHDCAD 133

RESULT 4  
US-09-984-186-12  
;; Sequence 12 Application US/09984186  
;; Patent No. US20020151011A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Fleet, Reinhard  
;; APPLICANT: Fournier, Alain  
;; APPLICANT: Guitton, Jean-Dominique  
;; APPLICANT: Jung, Gerard  
;; APPLICANT: Yeh, Patricia

;; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
;; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
;; CONTAINING SAID POLYPEPTIDES  
;; NUMBER OF SEQUENCES: 36  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
;; STREET: 500 Atcoia Road, 3C43  
;; CITY: Collegeville  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19426

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/984,186  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-Jan-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-Jul-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-Jan-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-Jan-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-984-186-12

Query Match 99.0%; Score 494; DB 10; Length 138;  
Best Local Similarity 98.8%; Pred. No. 8.6e-51;  
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNCRPN 60  
DB 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNCRPN 110  
QY 61 RRRPWCYVQVGLKPLVOECMVHDCAD 86  
DB 111 RRRPWCYVQVGLKPLVOECMVHDCAD 136

RESULT 5  
US-09-880-503-8  
;; Sequence 8 Application US/09880503  
;; Patent No. US20020131964A1  
;; GENERAL INFORMATION:  
;; APPLICANT: CINES, Douglas B  
;; APPLICANT: HIGAZI, Abd Al-Rouf  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
;; FILE REFERENCE: 9596-331  
;; CURRENT APPLICATION NUMBER: US/09/880,503  
;; CURRENT FILING DATE: 2001-06-13  
;; PRIOR APPLICATION NUMBER: US 60/212,847  
;; PRIOR FILING DATE: 2000-06-20  
;; NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-8

Query Match 99.0%; Score 494; DB 10; Length 143;  
Best Local Similarity 98.8%; Pred. No. 9e-51;  
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQOTYHAHRSNALQLGLGKHNCRPN 60  
|||||  
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQOTYHAHRSNALQLGLGKHNCRPN 107  
QY 61 RRRPWCYOVGLKPLVQECMVHDCAD 86  
|||||  
DB 108 RRRPWCYOVGLKPLVQECMVHDCAD 133

RESULT 6  
US-09-880-503-6

; Sequence 6, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 403  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-6

Query Match 99.0%; Score 494; DB 10; Length 403;  
Best Local Similarity 98.8%; Pred. No. 2.7e-50;  
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQOTYHAHRSNALQLGLGKHNCRPN 60  
|||||  
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQOTYHAHRSNALQLGLGKHNCRPN 107  
QY 61 RRRPWCYOVGLKPLVQECMVHDCAD 86  
|||||  
DB 108 RRRPWCYOVGLKPLVQECMVHDCAD 133

RESULT 7  
US-09-880-503-3

; Sequence 3, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 411

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-3

Query Match 99.0%; Score 494; DB 10; Length 411;  
Best Local Similarity 98.8%; Pred. No. 2.7e-50;  
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQOTYHAHRSNALQLGLGKHNCRPN 60  
|||||  
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQOTYHAHRSNALQLGLGKHNCRPN 107  
QY 61 RRRPWCYOVGLKPLVQECMVHDCAD 86  
|||||  
DB 108 RRRPWCYOVGLKPLVQECMVHDCAD 133

RESULT 8  
US-09-264-468B-1

; Sequence 1, Application US/09264468B  
; Patent No. US20020106775A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Jieyi  
; APPLICANT: Menabier, Vicki L.  
; APPLICANT: Henkin, Jack  
; APPLICANT: Smith, Richard A.  
; APPLICANT: Walter, Karl A.  
; APPLICANT: Severin, Jean M.  
; APPLICANT: Edalji, Rohinton  
; APPLICANT: Johnson Jr., Robert W.  
; APPLICANT: Holzman, Thomas F.  
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE  
; FILE REFERENCE: 6310 US P1  
; CURRENT APPLICATION NUMBER: US/09/264,468B  
; CURRENT FILING DATE: 1999-03-05  
; PRIOR APPLICATION NUMBER: US 09/036,361  
; PRIOR FILING DATE: 1998-03-06  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(20)  
; OTHER INFORMATION: Leader sequence  
; NAME/KEY: VARIANT  
; LOCATION: (279)...(279)  
; OTHER INFORMATION: Xaa = any amino acid  
; NAME/KEY: VARIANT  
; LOCATION: (302)...(302)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-264-468B-1

Query Match 99.0%; Score 494; DB 10; Length 431;  
Best Local Similarity 98.8%; Pred. No. 2.9e-50;  
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQOTYHAHRSNALQLGLGKHNCRPN 60  
|||||  
DB 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQOTYHAHRSNALQLGLGKHNCRPN 127  
QY 61 RRRPWCYOVGLKPLVQECMVHDCAD 86  
|||||  
DB 128 RRRPWCYOVGLKPLVQECMVHDCAD 153

RESULT 9  
US-09-084-491A-3

; Sequence 3, Application US/09084491A  
; Patent No. US20020061576A1  
; GENERAL INFORMATION:

```

; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: EBBERS, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/084,491A
; FILING DATE: 27-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,973
; REFERENCE/DOCKET NUMBER: PF318
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-084-491A-3

Query Match          43.3%; Score 216; DB 10; Length 372;
Best Local Similarity 47.0%; Pred. No. 6.2e-18;
Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

QY 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRPNDR 62
DB 25 CYEGNSAYRGTHSLTPSSGASCLPWNMILIGKVTYRQNPASQALGLGKHNCRPNDR 84
QY 63 RPPCYOVGKLPVQECMVHDC 85
DB 85 KPMCHVLKNNRLTWECYCDVPCS 107

RESULT 10
; US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-969-271-7

Query Match          43.3%; Score 216; DB 10; Length 562;
Best Local Similarity 47.0%; Pred. No. 9.7e-18;
```

```

Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

QY 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRPNDR 62
DB 215 CYEGNSAYRGTHSLTPSSGASCLPWNMILIGKVTYRQNPASQALGLGKHNCRPNDR 274
QY 63 RPPCYOVGKLPVQECMVHDC 85
DB 275 KPMCHVLKNNRLTWECYCDVPCS 297

RESULT 11
; US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hui-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Inocyte ID No. US20020156263A1 1001470CD1
; US-09-974-298-145

Query Match          43.1%; Score 215; DB 9; Length 562;
Best Local Similarity 45.9%; Pred. No. 1.3e-17;
Matches 39; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRPNDR 61
DB 126 TCYEDOGISYRGTWTAEAGACTWNNSSALAKYSGRRPDALRLGLGKHNCRPNDR 185
QY 62 RPPCYOVGKLPVQECMVHDC 86
DB 186 SKPMCYVFRAGKYSSEFCSTPACSE 210

RESULT 12
; US-09-912-559-3
; Sequence 3, Application US/09912559
; Patent No. US20020142316A1
; GENERAL INFORMATION:
; APPLICANT: ROEMISCH, JUERGEN
; APPLICANT: STOEHR, HANS-ARNOLD
; APPLICANT: FEUSNER, ANNETTE
; APPLICANT: LANG, WIEGAND
; APPLICANT: WEIMER, THOMAS
; APPLICANT: BECKER, MARGRET
; APPLICANT: MERLICH, CLAUDIA
; APPLICANT: MUTH-NUMANN, GUDRUN
; TITLE OF INVENTION: DETECTION OF THE FACTOR VII-ACTIVATING PROTEASE AND
; TITLE OF INVENTION: DETECTION METHODS USING SPECIFIC ANTIBODIES
; FILE REFERENCE: 06478.1457
; CURRENT APPLICATION NUMBER: US/09/912,559
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: DE 100 36 641.4
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: DE 100 50 040.4
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: DE 100 52 319.6
; PRIOR FILING DATE: 2000-10-21
; PRIOR APPLICATION NUMBER: DE 101 18 706.8
; PRIOR FILING DATE: 2001-04-12
```

NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 3  
LENGTH: 560  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-912-559-3

Query Match 37.8%; Score 188.5; DB 10; Length 560;  
Best Local Similarity 40.5%; Pred. No. 1.5e-14;  
Matches 34; Conservative 16; Mismatches 33; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASTDIMGRCPLPWNASATVLOQTYHAHRSNALQGLGKHNCRPNDR 62  
DB 194 CYVGDGYSYRGKMNRTVNHQACLYWNSHLLQENTNMFMEDAFETGIGEHNCRNPDAD 253  
QY 63 RPRWCYVQGLKPLVQ-CMVHDC 85  
DB 254 KPMCFIKYTNDRKWKWECVDSACS 277

RESULT 13  
US-09-912-559-4

Sequence 4, Application US/09912559  
Patent No. US20020142316A1

GENERAL INFORMATION:  
APPLICANT: ROEMISCH, JUERGEN  
APPLICANT: STOEHR, HANS-ARNOLD  
APPLICANT: FEUSSNER, ANNETTE  
APPLICANT: LANG, WIEGAND  
APPLICANT: WEIMER, THOMAS  
APPLICANT: BECKER, MARGRET  
APPLICANT: NERLICH, CLAUDIA  
APPLICANT: MUTH-NAUMANN, GUDRUN  
TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND  
FILE REFERENCE: 06478.1457  
CURRENT FILING DATE: 2001-07-26  
CURRENT APPLICATION NUMBER: US/09/912.559  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: DE 100 36 641.4  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: DE 100 50 040.4  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: DE 100 52 319.6  
PRIOR FILING DATE: 2000-10-21  
PRIOR APPLICATION NUMBER: DE 101 18 706.8  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 560  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-912-559-4

Query Match 37.8%; Score 188.5; DB 10; Length 560;  
Best Local Similarity 40.5%; Pred. No. 1.5e-14;  
Matches 34; Conservative 16; Mismatches 33; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASTDIMGRCPLPWNASATVLOQTYHAHRSNALQGLGKHNCRPNDR 62  
DB 194 CYVGDGYSYRGKMNRTVNHQACLYWNSHLLQENTNMFMEDAFETGIGEHNCRNPDAD 253  
QY 63 RPRWCYVQGLKPLVQ-CMVHDC 85  
DB 254 KPMCFIKYTNDRKWKWECVDSACS 277

RESULT 14  
US-09-761-120-42  
Sequence 42, Application US/09761120  
Patent No. US20020037847A1  
GENERAL INFORMATION:

APPLICANT: Folkman, M. Judah  
APPLICANT: O'Reilly, Michael  
TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin  
FILE REFERENCE: 05940-0151 (43171-252068)  
CURRENT FILING DATE: 2001-01-16  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 09/309,821  
PRIOR FILING DATE: 1997-05-30  
NUMBER OF SEQ ID NOS: 47  
SOFTWARE: Patentln version 3.0  
SEQ ID NO 42  
LENGTH: 368  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Kringle 1-4 BKLS  
US-09-761-120-42

Query Match 30.3%; Score 151; DB 10; Length 368;  
Best Local Similarity 38.2%; Pred. No. 2.3e-10;  
Matches 34; Conservative 11; Mismatches 30; Indels 14; Gaps 5;

QY 3 CYEGNGHFYRGKASTDIMGRCPLPWNASATVLOQTYHAHRSNALQGLGKHNCRPNDR 59  
DB 275 CYHGDGOSYRGKSTSTTTGKKCOSWS--MTPRHQKTPENYPNAGL-TMYNCRNPD 328  
QY 60 NRRRWCYVQGLKPLV--OECMVHDC 86  
DB 329 ADKGPWCFT--TDPSTWMEYCNKKCS 354

RESULT 15  
US-09-873-676-1  
Sequence 1, Application US/09873676  
Patent No. US2002007289A1  
GENERAL INFORMATION:  
APPLICANT: MacDonald, Nicholas J.  
APPLICANT: Sim, Kim L.  
TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use  
FILE REFERENCE: 05213-0378 (43170-259333)  
CURRENT FILING DATE: 2001-06-04  
CURRENT APPLICATION NUMBER: US/09/873.676  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: US 60/209,065  
PRIOR FILING DATE: 2001-05-08  
PRIOR APPLICATION NUMBER: US 60/289,387  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: Patentln version 3.1  
SEQ ID NO 1  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-873-676-1

Query Match 29.9%; Score 149; DB 10; Length 378;  
Best Local Similarity 38.6%; Pred. No. 4.1e-10;  
Matches 34; Conservative 10; Mismatches 30; Indels 14; Gaps 5;

QY 3 CYEGNGHFYRGKASTDIMGRCPLPWNASATVLOQTYHAHRSNALQGLGKHNCRPNDR 59  
DB 285 CYHGDGOSYRGKSTSTTTGKKCOSWS--MTPRHQKTPENYPNAGL-TMYNCRNPD 338  
QY 60 NRRRWCYVQGLKPLV--OECMVHDC 85  
DB 339 ADKGPWCFT--TDPSTWMEYCNKKCS 363

Search completed: November 11, 2002, 13:03:15  
Job time: 15 secs

100



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 12:57:46 ; Search time 38 Seconds  
(without alignments)  
217.567 Million cell updates/sec

Title: PCT-US02-27855-1

Perfect score: 499  
Sequence: 1 KTCYEGNGHFYRGRASIDYM.....YVQVGLKPLVQECMVHDCAD 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR73:\*

1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	494	99.0	431	1 UKHR	u-plasminogen acti
2	459	92.0	433	1 UKBAY	u-plasminogen acti
3	409	82.0	442	1 UKPG	u-plasminogen acti
4	400	80.2	432	1 S18932	u-plasminogen acti
5	383	76.8	433	1 JN0560	u-plasminogen acti
6	381	76.4	433	1 UKMS	u-plasminogen acti
7	336	47.3	431	2 JS0599	u-plasminogen acti
8	236	47.3	477	1 A34369	u-plasminogen acti
9	236	47.3	477	2 JS0598	u-plasminogen acti
10	226	45.3	559	1 A35029	u-plasminogen acti
11	218	43.7	394	2 JS0600	u-plasminogen acti
12	216	43.3	562	1 UKHUT	u-plasminogen acti
13	215	43.1	291	2 I38098	u-plasminogen acti
14	215	43.1	559	1 A29941	u-plasminogen acti
15	208	41.7	477	2 JS0597	u-plasminogen acti
16	204.5	41.0	558	1 JS0578	u-plasminogen acti
17	203.5	40.8	434	1 A35005	u-plasminogen acti
18	193	38.7	555	1 A46688	hepatocyte growth
19	188.5	37.8	560	1 JC4795	plasma hyaluronan
20	179.5	36.0	603	2 S28941	coagulation factor
21	170.5	34.2	615	1 KFH012	coagulation factor
22	158	31.7	89	2 A60140	plasmin (EC 3.4.21
23	156	31.3	120	2 B61545	plasmin (EC 3.4.21
24	155.5	31.3	1420	2 A32869	apolipoprotein(a)
25	155.5	31.2	593	2 S45281	coagulation factor
26	155	31.1	4548	1 S00657	apoptosis factor
27	154	30.9	812	1 PLBO	plasmin (EC 3.4.21
28	153	30.7	790	1 PLPG	plasmin (EC 3.4.21
29	152.5	30.6	169	2 A40522	plasmin (EC 3.4.21

## ALIGNMENTS

## RESULT 1

UKHR  
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human  
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen  
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen  
in form  
C:Species: Homo sapiens (man)  
C:Date: 17-Dec-1982 #sequence\_revision 04-Dec-1986 #text\_change 15-Sep-2000  
C:Accession: A00931; I52209; J0102; A37561; I38102; S65783; A37562; A37563; A37564;  
R:Ricciolo, A.; Grimaldi, G.; Verde, P.; Sebastiao, G.; Boast, S.; Blasl, F.  
Nucleic Acids Res. 13, 2759-2771, 1985  
A:Title: The human urokinase-plasminogen activator gene and its promoter.  
A:Reference number: A00931; MUID:85215647; PMID:2967867  
A:Accession: A00931  
A:Molecule type: DNA  
A:Residues: 1-431 <R1C>  
A:Cross-references: GB:X02419; NID:937601; PIDN:CAA26268.1; PID:9340175  
A:Note: The authors translated the codon ATG for residue 214 as Ile  
R:Nagamine, Y.; Pearson, D.; Gratian, M.  
Biochem. Biophys. Res. Commun. 132, 563-569, 1985  
A:Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porci  
A:Reference number: I52209; MUID:86050639; PMID:3935305  
A:Accession: I52209  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 145-161 <NAG1>  
A:Cross-references: GB:K03027; NID:9340174; PIDN:AAA61257.1; PID:9340175  
R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyam  
Gene 36, 183-188, 1985  
A:Title: Molecular cloning of cDNA coding for human preprourokinase.  
A:Reference number: J0102; MUID:86056954; PMID:2415429  
A:Accession: J0102  
A:Molecule type: mRNA  
A:Residues: 1-213, 'I', 215-431 <NAG2>  
A:Cross-references: GB:K03226; NID:9340155; PIDN:AA097138.1; PID:9340158; GB:D00244;  
R:Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasl, F.  
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984  
A:Title: Identification and primary sequence of an unspliced human urokinase poly(A)+  
A:Reference number: A37561; MUID:84272706; PMID:5589620  
A:Accession: A37561  
A:Molecule type: mRNA  
A:Residues: 66-431 <VER>  
A:Cross-references: GB:D00244; NID:9220138  
R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Coliau, B.; Chuchana, P.; van Els  
DNA 4, 139-146, 1985  
A:Title: Molecular cloning, sequencing, and expression in Escherichia coli of human p  
A:Reference number: I38102; MUID:85203559; PMID:3888571  
A:Accession: I38102  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>  
A:Cross-references: EMBL:X02760; NID:935297; PIDN:CAA26535.1; PID:935298

R.Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, K.; Bloch, M. *Biophys. Acta* 1293, 83-89, 1996  
 A>Title: Characterization of single chain urokinase-type plasminogen activator with a no  
 A:Accession: S65783; MUID:96186279; PMID:8652631  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 21-140; 'L', 142-213; 'I', 215-431 <YOS>  
 A:Cross-references: EMBL:D11143; NID:93131467; PIDN:BA01919.1; PID:q1199928  
 R.Gunzer, M.A.; Steffens, G.O.; Gellings, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982  
 A>Title: The primary structure of high molecular mass urokinase from human urine.  
 A:Reference number: A37562; MUID:83055084; PMID:6754569  
 A:Accession: A37562  
 A:Molecule type: protein  
 A:Residues: 21-177 <GUN>  
 R.Schaller, J.; Nick, H.; Rickli, E.E.; Gillissen, D.; Lergier, W.; Studer, R.O.  
 Eur. J. Biochem. 123, 251-257, 1982  
 A>Title: Human low-molecular-weight urinary urokinase. Partial characterization and pre  
 A:Reference number: A37563; MUID:83003608; PMID:6749491  
 A:Accession: A37563  
 A:Molecule type: protein  
 A:Residues: 156-176; 179-193; 'T', 195; 'T', 197-224 <SCH>  
 R.Steffens, G.J.; Gunzer, M.A.; Ottling, F.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982  
 A>Title: The complete amino acid sequence of low molecular mass urokinase from human ur  
 A:Reference number: A37564; MUID:83055099; PMID:6754572  
 A:Accession: A37564  
 A:Molecule type: protein  
 A:Residues: 158-410 <SRE>  
 R.Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.  
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990  
 A>Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan  
 A:Reference number: A35683; MUID:90365737; PMID:2393398  
 A:Accession: A35683  
 A:Molecule type: protein  
 A:Residues: 21-30; 'X', 32; 'X', 34-38; 'X', 40-43 <KEN>  
 R.Rabdan, S.A.; Desjardins, J.; Bell, A.W.; Banyville, D.; Mazur, A.; Henkin, J.; Goltz  
 Bloch, M. *Biophys. Res. Commun.* 173, 1058-1064, 1990  
 A>Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li  
 A:Reference number: A36697; MUID:91097529; PMID:2125213  
 A:Accession: A36697  
 A:Molecule type: protein  
 A:Residues: 21-34 <RAB>  
 R.Li, X.; Bokman, A.M.; Ilines, M.; Smith, R.A.G.; Dobson, C.M.  
 submitted to the Brookhaven Protein Data Bank, July 1993  
 A:Reference number: A51255; PDB:1KDU  
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R.Li, X.; Smith, R.A.G.; Dobson, C.M.  
 Biochemistry 31, 9562-9571, 1992  
 A>Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain  
 A:Reference number: A44375; MUID:93003110; PMID:1327116  
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
 R.Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,  
 submitted to the Brookhaven Protein Data Bank, January 1994  
 A:Reference number: A66822; PDB:1UOK  
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R.Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;  
 submitted to the Brookhaven Protein Data Bank, July 1995  
 A:Reference number: A6058; PDB:1LLW  
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426  
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A  
 C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, B  
 C:Genetics: GDB:PLAU  
 A:Cross-references: GDB:119497; OMIM:191840  
 A:Map position: 10q24-10q24  
 A:Introns: 19/1; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3  
 C:Function:  
 A:Description: proteolytically activates plasminogen  
 A:Pathway: fibrinolysis  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine prote  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-431/Product: urokinase-type plasminogen activator, single chain form #status pre  
 F:21-177/Product: urokinase-type plasminogen activator chain A #status experimental <  
 F:31-62/Domain: EGF homology <EGF>  
 F:70-151/Domain: kringle homology <KR>  
 F:156-177/Product: urokinase-type plasminogen activator chain A #status experimental  
 F:179-431/Product: urokinase-type plasminogen activator chain B #status experimental  
 F:179-431/Domain: trypsin homology <TRY>  
 F:31-39; 33-51; 53-62; 70-151; 91-133; 122-146; 168-299; 209-225; 217-288; 313-382; 345-361; 372  
 F:38/Binding site: carboxylate (Thr) (covalent) #status predicted  
 F:178-179/Cleavage site: lys-ile (plasmin) #status experimental  
 F:224,275; 376/Active site: His, Asp, Ser (covalent) #status experimental  
 F:322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 99.0%; Score 494; DB 1; Length 431;  
 Best Local Similarity 98.8%; Pred. No 4e-47; 0; Indels 0; Gaps 0;  
 Matches 85; Conservative 1; Mismatches 0;  
 QY 1 KTCYEGNGHFYRGKASTDTGRCPLPMSATVLOQTYHAHRSNALDGLGKHNYCRNPDN 60  
 DB 68 KTCYEGNGHFYRGKASTDTGRCPLPMSATVLOQTYHAHRSNALDGLGKHNYCRNPDN 127  
 QY 61 RRRPWCYVQVGLKPLVQECVYHDCAD 86  
 DB 128 RRRPWCYVQVGLKPLVQECVYHDCAD 153

## RESULT 2

UKBAY  
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon  
 C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)  
 C:Date: 31-Dec-1991 #sequence, revision 31-Dec-1991 #text\_change 18-Jun-1999  
 A:Accession: S14687; S08651  
 R:Au, Y.P.T.; Wang, T.W.; Clowes, A.W.  
 Nucleic Acids Res. 18, 3411, 1990  
 A>Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasmin  
 A:Reference number: S14687; MUID:90287734; PMID:2113276  
 A:Accession: S14687  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <ANDY>  
 A:Cross-references: EMBL:X51935; NID:938130; PIDN:CA36200.1; PID:938131  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology;  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine protease  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-176/Product: plasminogen activator chain A #status predicted <ACH>  
 F:30-61/Domain: EGF homology <EGF>  
 F:69-150/Domain: kringle homology <KR>  
 F:178-431/Product: plasminogen activator chain B #status predicted <BCH>  
 F:178-431/Domain: trypsin homology <TRY>  
 F:167-298; 208-224; 216-287; 315-384; 347-363; 374-402/Disulfide bonds: #status predicted  
 F:223,274,378/Active site: His, Asp, Ser #status predicted  
 F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 92.0%; Score 459; DB 1; Length 433;  
 Best Local Similarity 93.0%; Pred. No. 3.1e-43;  
 Matches 80; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTGRCPLPMSATVLOQTYHAHRSNALDGLGKHNYCRNPDN 60  
 DB 67 KTCYEGNGHFYRGKASTDTGRCPLPMSATVLOQTYHAHRSNALDGLGKHNYCRNPDN 126  
 QY 61 RRRPWCYVQVGLKPLVQECVYHDCAD 86  
 DB 127 RRRPWCYVQVGLKPLVQECVYHDCAD 152

## RESULT 3

UKBE  
 u-plasminogen activator (EC 3.4.21.73) precursor - pig  
 N:Alternate names: uPA  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 04-Dec-1986 #sequence, revision 17-Mar-1987 #text\_change 07-Aug-1998

C:Accession: A00932  
 R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.  
 Nucleic Acids Res. 12, 9525-9541, 1984  
 A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.  
 A:Reference number: A00932; MUID:85087954; PMID:6096832  
 A:Accession: A00932  
 A:Molecule type: DNA  
 A:Residues: 1-240, 'H', 242-442 <NAG1>  
 A:Experimental source: kidney cell line LLC-PK1  
 R:Nagamine, Y.  
 submitted to the Protein Sequence Database, December 1986  
 A:Reference number: A37566  
 A:Contents: annotation; correction to residue 241  
 C:Genetics:  
 A:Insertions: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3  
 A:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F:21-64/Domain: EGF homology <EGF>  
 F:72-153/Domain: kringle homology <KR>  
 F:190-430/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
 F:152/Binding site: carbonyl (Asn) (covalent) #status predicted  
 F:179-310-220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted  
 F:235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 82.0%; Score 409; DB 1; Length 442;  
 Best Local Similarity 82.4%; Pred. No. 1.1e-37;  
 Matches 70; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGRKASPTMGRCPLPWNATVLTQOTYHAHRSNALDGLGKHNCRPN 60  
 DB 70 QTCYEGNGHYRGRKASPTMGRCPLPWNATVLTQOTYHAHRSNALDGLGKHNCRPN 129  
 QY 61 RRRPWCYVQGLKPLVQECMVHDC 85  
 DB 130 QRRPWCYVQGLKPLVQECMVHDC 154

RESULT 4  
 S18932  
 u-plasminogen activator (EC 3.4.21.73) precursor - rat  
 N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 18-Oct-1989 #sequence\_revision 10-Feb-1995 #text\_change 18-Jun-1999  
 C:Accession: S24604; I60186; I53472; S18932  
 R:Rabban, S.A.  
 submitted to the EMBL Data Library, April 1992  
 A:Reference number: S24604  
 A:Accession: S24604  
 A:Molecule type: mRNA  
 A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>  
 A:Cross-references: EMBL:X65651; NID:957456; PIDN:CAA6601.1; PID:957457  
 A:Experimental source: tissue kidney  
 R:Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.  
 Cancer Res. 52, 2489-2496, 1992  
 A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act  
 A:Reference number: I60186; MUID:92233409; PMID:1568219  
 A:Accession: I60186  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-432 <RES>  
 A:Cross-references: EMBL:X63434; NID:957465; PIDN:CAA5028.1; PID:957466  
 A:Experimental source: Strain Fischer 344; tissue mammary  
 R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blas, F.; Rossi, G.  
 FEBS Lett. 306, 193-198, 1992  
 A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated in  
 A:Reference number: I53472; MUID:92339549; PMID:1321734  
 A:Accession: I53472  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 31-62 <RE2>

A:Cross-references: EMBL:X66907; NID:9396200; PIDN:CAA47356.1; PID:9398279  
 C:Genetics:  
 A:Gene: uPA  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology;  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F:31-62/Domain: EGF homology <EGF>  
 F:70-151/Domain: kringle homology <KR>  
 F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BC  
 F:179-430/Domain: trypsin homology <TRY>  
 F:168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted  
 F:225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 80.2%; Score 400; DB 1; Length 432;  
 Best Local Similarity 80.0%; Pred. No. 1.1e-36;  
 Matches 68; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGRKASPTMGRCPLPWNATVLTQOTYHAHRSNALDGLGKHNCRPN 60  
 DB 68 KTCYEGNGHYRGRKASPTMGRCPLPWNATVLTQOTYHAHRSNALDGLGKHNCRPN 127  
 QY 61 RRRPWCYVQGLKPLVQECMVHDC 85  
 DB 128 QRRPWCYVQGLKPLVQECMVHDC 152

RESULT 5  
 JN0560  
 u-plasminogen activator (EC 3.4.21.73) precursor - bovine

N:Alternate names: uPA  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C:Accession: JN0560  
 R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.  
 Gene 125, 177-183, 1993  
 A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and in  
 A:Reference number: JN0560; MUID:93216119; PMID:8385052  
 A:Accession: JN0560  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <KRA>

A:Cross-references: GB:I03546; NID:9163800; PIDN:AA51419.1; PID:9163801  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology;  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-179/Product: plasminogen activator chain A #status predicted <A1>  
 F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F:33-64/Domain: EGF homology <EGF>  
 F:72-153/Domain: kringle homology <KR>  
 F:181-423/Product: plasminogen activator chain B #status predicted <A2>  
 F:181-421/Domain: trypsin homology <TRY>  
 F:170-301,211-227,219-290,347-363,374-402/Disulfide bonds: #status predicted  
 F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 76.8%; Score 383; DB 1; Length 433;  
 Best Local Similarity 75.3%; Pred. No. 8.4e-35;  
 Matches 64; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGRKASPTMGRCPLPWNATVLTQOTYHAHRSNALDGLGKHNCRPN 60  
 DB 70 KTCYEGNGHYRGRKASPTMGRCPLPWNATVLTQOTYHAHRSNALDGLGKHNCRPN 129  
 QY 61 RRRPWCYVQGLKPLVQECMVHDC 85  
 DB 130 QRRPWCYVQGLKPLVQECMVHDC 154

RESULT 6  
 UKMS  
 u-plasminogen activator (EC 3.4.21.73) precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 18-Jun-1999  
 C:Accession: A29420; A24615

R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.  
 Biochemistry 26, 8270-8279, 1987  
 A:Title: The murine urokinase-type plasminogen activator gene.  
 A:Reference number: A29420; MUID:88163489; PMID:2831940  
 A:Accession: A29420  
 A:Molecule type: DNA  
 A:Residues: 1-433 <DEL>  
 A:Cross-references: GB:M17922; NID:9202296; PIDN:AAA0539.1; PID:9202297  
 R;Bell, D.; Vassalli, J.D.; Compele, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, E.  
 J. Biochem. 148, 225-232, 1985  
 A:Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase  
 A:Reference number: A24615; MUID:85179474; PMID:2985383  
 A:Accession: A24615  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <DEL>  
 A:Cross-references: GB:X02389; NID:955127; PIDN:CAA26231.1; PID:955128  
 C:Gene: 19/3: 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F:32-63/Domain: EGF homology <EGF>  
 F:71-152/Domain: kringle homology <KR>  
 F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>  
 F:169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted  
 F:226,277,378/Active site: His, Asp, Ser #status predicted  
 Query Match 75.4%; Score 381; DB 1; Length 433;  
 Best Local Similarity 72.9%; Pred. No. 1,4e-34;  
 Matches 62; Conservative 11; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 KTYEGNGHFYRGKASTDTMGRCPLPWSATYVLOOTYHAHRSNALQGLGKHNCRNDN 60  
 DB 69 KTYEGNGSYRGKANTDTKGRPCPLMNAVALQKRYNAHRDASLIGKHNCRNDN 128  
 QY 61 RRPWCYVQVGLKPLVQECMVHDC 85  
 DB 129 QKRWPCYVQIGLRFVQECMVHDCS 153  
 RESULT 7  
 JS0599  
 t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat  
 N:Alternate names: tissue plasminogen activator  
 C:Species: Desmodus rotundus (common vampire bat)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C:Accession: JS0599  
 R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boldol, W.; Brinmann, P.; Alagon, A.; Dor  
 Gene 105, 229-237, 1991  
 A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
 A:Reference number: JS0597; MUID:92039036; PMID:1937019  
 A:Accession: JS0599  
 A:Molecule type: mRNA  
 A:Residues: 1-431 <KRA>  
 A:Cross-references: GB:M63989; NID:9166076; PIDN:AAA31594.1; PID:9166077  
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-36/Domain: propeptide #status predicted <PRO>  
 F:37-431/Product: plasminogen activator beta #status predicted <PLA>  
 F:81-74/Domain: EGF homology <EGF>  
 F:82-163/Domain: kringle homology <KR>  
 F:180-425/Domain: trypsin homology <TRY>  
 F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bon  
 F:139,352/binding site: carboxylate (Asn) (covalent) #status predicted  
 F:179-180/Cleavage site: His-Ser (Plasmin) #status predicted  
 F:226,275,382/Active site: His, Asp, Ser #status predicted  
 F:345-361,378-406/Disulfide bonds: #status predicted  
 Query Match 47.3%; Score 236; DB 2; Length 431;  
 Best Local Similarity 48.8%; Pred. No. 1.7e-18;

Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;  
 QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATYVLOOTYHAHRSNALQGLGKHNCRNDN 61  
 DB 81 TCYDQGYTYRGTWSTSSGAQCINMNSNLTFRYNGRRSDAITGLGNHNCRNPDN 140  
 QY 62 RRPWCYVQVGLKPLVQECMVHDC 85  
 DB 141 SKPWCYVAKSKFIEFCSPVCS 164  
 RESULT 8  
 A3369  
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)  
 C:Species: Megaderma lyra  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A3369  
 R:Gardell, S.J.; Duong, L.T.; Dielh, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Ja  
 J. Biol. Chem. 264, 17947-17952, 1989  
 A:Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plas  
 A:Reference number: A3369; MUID:90036867; PMID:2509450  
 A:Accession: A3369  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-477 <GAR>  
 A:Cross-references: GB:J05082; NID:9166080; PIDN:AAA31596.1; PID:9166081  
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat  
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-36/Domain: propeptide #status predicted <PRO>  
 F:37-477/Product: plasminogen activator #status predicted <PLA>  
 F:42-79/Domain: EGF homology <EGF>  
 F:87-120/Domain: kringle homology <KR>  
 F:128-209/Domain: trypsin homology <TRY>  
 F:226-471/Domain: trypsin homology <TRY>  
 F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,35  
 F:272,321,428/Active site: His, Asp, Ser #status predicted  
 Query Match 47.3%; Score 236; DB 1; Length 477;  
 Best Local Similarity 48.8%; Pred. No. 1.9e-18;  
 Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;  
 QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATYVLOOTYHAHRSNALQGLGKHNCRNDN 61  
 DB 127 TCYDQGYTYRGTWSTSSGAQCINMNSNLTFRYNGRRSDAITGLGNHNCRNPDN 186  
 QY 62 RRPWCYVQVGLKPLVQECMVHDC 85  
 DB 187 SKPWCYVAKSKFIEFCSPVCS 210  
 RESULT 9  
 JS0598  
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat  
 N:Alternate names: tissue plasminogen activator  
 C:Species: Desmodus rotundus (common vampire bat)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C:Accession: JS0598  
 R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boldol, W.; Brinmann, P.; Alagon, A.;  
 Gene 105, 229-237, 1991  
 A:Title: The plasminogen activator family from the salivary gland of the vampire bat  
 A:Reference number: JS0597; MUID:92039036; PMID:1937019  
 A:Accession: JS0598  
 A:Molecule type: mRNA  
 A:Residues: 1-477 <KRA>  
 A:Cross-references: GB:M63988; NID:9166074; PIDN:AAA31593.1; PID:9166075  
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat  
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-36/Domain: propeptide #status predicted <PRO>  
 F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>  
 F:42-79/Domain: trypsin homology <TRY>  
 F:87-120/Domain: EGF homology <EGF>

F:128-209/Domain: kringle homology <KRG>  
 F:226-471/Domain: trypsin homology <TRY>  
 F:42-70-79,87-96,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4  
 F:185,396/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:225-226/Cleavage site: His-Ser (plasmin) #status predicted  
 F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 47.3%; Score 236; DB 2; Length 477;  
 Best Local Similarity 48.8%; Pred. No. 1.9e-18;  
 Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRCPLPWNSTAVYLOOTYHAHRSNALQGLGKHNHCNPNDR 61  
 DB 127 TCKKDGQVYTRKGTWSTSESGACINMNSNLIRRTNGRSDAITLGLGNHNYCNRPNDR 186

QY 62 RRPWCYVQVGLKPLVQECVHDC 85  
 DB 187 SKPWCYVIRAKRFTSESCVPVCS 210

## RESULT 10

A35029  
 t-plasminogen activator (EC 3.4.21.68) precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A35029; A31597  
 R:Feng, P.; Ohlsson, M.; Ny, T.  
 J. Biol. Chem. 265, 2022-2027, 1990

A:Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec  
 A:Reference number: A35029; MUID:90130448; PMID:2105315  
 A:Accession: A35029

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-559 <REN>

A:Cross-references: GB:M31197; NID:9207429; PIDN:AAA42261.1; PID:9207431; GB:J05226

R:Ny, T.; Leonardson, G.; Hsueh, A.-J.W.  
 DNA 7, 671-677, 1988

A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator  
 A:Reference number: A31597; MUID:89170114; PMID:3148445

A:Accession: A31597

A:Molecule type: mRNA

A:Residues: 1-379, 'K', 381-559 <NYT>

A:Cross-references: GB:M23697; NID:9530159; PIDN:AAA41812.1; PID:9530160

C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-29/Domain: propeptide #status predicted <PRO>

F:30-559/Product: t-plasminogen activator #status predicted <MAT>

F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>

F:38-75/Domain: fibronectin type I repeat homology <FI1>

F:93-116/Domain: EGF homology <EGF>

F:124-205/Domain: kringle homology <KR1>

F:213-294/Domain: kringle homology <KR2>

F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>

F:309-553/Domain: trypsin homology <TRY>

F:38-66,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4

F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:308-309/Cleavage site: Arg-Tile (plasmin, trypsin) #status predicted

F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 45.3%; Score 226; DB 1; Length 559;  
 Best Local Similarity 49.4%; Pred. No. 2.8e-17;  
 Matches 41; Conservative 9; Mismatches 33; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRCPLPWNSTAVYLOOTYHAHRSNALQGLGKHNHCNPNDR 61

DB 123 TCYEGNGHYRGKASTDTMGRCPLPWNSTAVYLOOTYHAHRSNALQGLGKHNHCNPNDR 182

QY 62 RRPWCYVQVGLKPLVQECVHDC 84

DB 183 VKPWCYVIRAKRFTSESCVPVCS 205

## RESULT 11

JS0600

t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat

N:Alternate names: tissue plasminogen activator

C:Species: Desmodus rotundus (common vampire bat)

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999

C:Accession: JS0600

R:Kreitschmar, J.; Haendler, B.; Langer, G.; Boldol, W.; Bräunmann, P.; Alagon, A.;

Gene 105, 229-237, 1991

A:Title: The plasminogen activator family from the salivary gland of the vampire bat

A:Reference number: JS0597; MUID:92039036; PMID:1937019

A:Accession: JS0600

A:Molecule type: mRNA

A:Residues: 1-394 <KRA>

A:Cross-references: GB:M63990; NID:9166078; PIDN:AAA31595.1; PID:9166079

A:Note: The authors translated the codon ATC for residue 75 as Thr

C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-36/Domain: propeptide #status predicted <PRO>

F:37-394/Product: plasminogen activator gamma #status predicted <PLA>

F:45-126/Domain: trypsin homology <TRY>

F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bond

F:149-143/Cleavage site: His-Ser (plasmin) #status predicted

F:189,238,345/Active site: His, Asp, Ser #status predicted

F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.7%; Score 218; DB 2; Length 394;  
 Best Local Similarity 45.2%; Pred. No. 1.6e-16;  
 Matches 38; Conservative 12; Mismatches 34; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRCPLPWNSTAVYLOOTYHAHRSNALQGLGKHNHCNPNDR 61

DB 44 TCKKDGQVYTRKGTWSTSESGACINMNSNLIRRTNGRSDAITLGLGNHNYCNRPNDR 103

QY 62 RRPWCYVQVGLKPLVQECVHDC 85

DB 104 SKPWCYVIRAKRFTSESCVPVCS 127

## RESULT 12

UKR007

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human

N:Alternate names: t-PA; tissue plasminogen activator

C:Species: Homo sapiens (man)

C>Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000

C:Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645;

R:Ny, T.; Elgh, F.; Lund, B.

Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984

A:Title: The structure of the human tissue-type plasminogen activator gene: correlati

A:Reference number: A94004; MUID:84298137; PMID:6089198

A:Accession: A94004

A:Molecule type: DNA

A:Residues: 1-562 <NYT>

A:Cross-references: GB:I00141

A:Note: The codon given for residue 93 (ACC) is inconsistent with the authors' transla

R:Frietzer Degen, S.J.; Rajput, B.; Reich, E.

J. Biol. Chem. 261, 6972-6985, 1986

A:Title: The human tissue plasminogen activator gene.

A:Reference number: A23529; MUID:86196143; PMID:3009482

A:Accession: A23529

A:Molecule type: DNA

A:Residues: 1-562 <DEG>

A:Cross-references: GB:K03021; NID:9339817; PIDN:AAA98809.1; PID:9339818

R:Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.

Agric. Biol. Chem. 55, 1225-1232, 1991

A:Title: Purification and characterization of tissue plasminogen activator secreted b

A:Reference number: J0562; MUID:91291340; PMID:1368681

A:Accession: J0562

A:Molecule type: mRNA

A:Residues: 31-562 <TTA>

A:Cross-references: DBU:J01096; NID:9220128; PIDN:BA00881.1; PID:9441174

A:Experimental source: embryonic lung fibroblast IMR-90 cells  
 A:Note: part of this sequence, including the amino end of the mature protein, was confir  
 R:Penicka, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett  
 Nature 301, 214-221, 1983  
 A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche  
 A:Reference number: A93593; MUID:83115262; PMID:6337343  
 A:Accession: A93593  
 A:Molecule type: mRNA  
 A:Residues: 1-562 <PEN>  
 A:Cross-references: GB:I00141  
 A:Experimental source: melanoma cells  
 R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.  
 Nucleic Acids Res. 16, 5695, 1988  
 A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe  
 A:Reference number: S03125; MUID:88262579; PMID:3133640  
 A:Accession: S03125  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-562 <SAS>  
 A:Cross-references: EMBL:X07393; NID:g37243; PIDN:CAA30302.1; PID:g37244  
 A:Experimental source: fetal lung cells  
 R:Kagilant, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma  
 FEBS Lett. 189, 145-149, 1985  
 A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen  
 A:Reference number: A91343; MUID:85285620; PMID:3896853  
 A:Accession: A91343  
 A:Molecule type: mRNA  
 A:Residues: 1-36, 86-433, 'E', 435-562 <KAG>  
 A:Experimental source: Detroit 582 cells; ATCC 138  
 R:Edlund, T.; Ny, T.; Randy, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983  
 A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen acti  
 A:Reference number: A93951; MUID:83169656; PMID:6572897  
 A:Accession: A93951  
 A:Molecule type: mRNA  
 A:Residues: 251-358 <EDL>  
 A:Experimental source: melanoma cells  
 R:Polh, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.  
 Biochemistry 23, 3701-3707, 1984  
 A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived an  
 differences.  
 A:Reference number: A90488; MUID:85000468; PMID:6433976  
 A:Contents: annotation; melanoma cells; partial sequence of residues 36-562, active and  
 R:Polh, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.  
 FEBS Lett. 168, 29-32, 1984  
 A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator.  
 A:Reference number: A91322; MUID:84158956; PMID:6538514  
 A:Accession: A91322  
 A:Molecule type: protein  
 A:Residues: 33-45;311-320 <POH>  
 A:Experimental source: uterus  
 A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln  
 R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.  
 J. Biol. Chem. 261, 14214-14218, 1986  
 A:Reference number: A57567; MUID:87033611; PMID:3021732  
 A:Contents: annotation; fibrin binding site  
 R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger  
 EMBO J. 5, 3525-3530, 1986  
 A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac  
 A:Reference number: A37568; MUID:87161761; PMID:3030730  
 A:Contents: annotation; fibrin binding site  
 R:Dodd, I.; Nunn, B.; Robinson, J.H.  
 Thromb. Haemost. 59, 523-528, 1988  
 A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type p  
 A:Reference number: A60902; MUID:89044681; PMID:3142086  
 A:Contents: annotation; novel forms of expressed recombinant t-PA  
 R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.;  
 Mol. Biol. Med. 3, 279-292, 1986  
 A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its exp  
 A:Reference number: A54645; MUID:86284200; PMID:3050401  
 A:Accession: A54645  
 A:Molecule type: mRNA  
 A:Residues: 1-562 <HAN>

A:Cross-references: GB:M15518; NID:g190031; PIDN:AAA60111.1; PID:g190032  
 A:Note: parts of this sequence were confirmed by peptide sequencing  
 R:Reddy, V.B.; Garramone, A.J.; Sasaki, H.; Wei, C.  
 DNA 6, 461-472, 1987  
 A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells  
 A:Reference number: I60110; MUID:88054470; PMID:2824147  
 A:Accession: I60110  
 A:Status: translated from GB/EMBL/DBD  
 A:Molecule type: mRNA  
 A:Residues: 1-562 <RES>  
 A:Cross-references: GB:M18182; NID:g340176; PIDN:AAA36800.1; PID:g340177  
 R:Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.  
 J. Biol. Chem. 260, 11223-11230, 1985  
 A:Title: Isolation and characterization of the human tissue-type plasminogen activato  
 A:Reference number: I55232; MUID:85289338; PMID:3161893  
 A:Accession: I55232  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-36 <RED>  
 A:Cross-references: GB:M1890; NID:g339837; PIDN:AAA61213.1; PID:g339839  
 C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a sing  
 C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond.  
 C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repea  
 C:Genetics:  
 A:Gene: GDB:PLAT  
 A:Cross-references: GDB:119496; OMIM:173370  
 A:Map position: 8p12-8p12  
 A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3;  
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat  
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase  
 F:1-33/Domain: signal sequence #status predicted <SIG>  
 F:24-32/Domain: propeptide #status predicted <PRO>  
 F:33-56/Product: t-plasminogen activator #status experimental <MAT>  
 F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>  
 F:41-78/Domain: fibronectin type I repeat homology <IF1>  
 F:86-119/Domain: EGF homology <EGF>  
 F:127-208/Domain: kringle homology <KR1>  
 F:215-296/Domain: kringle homology <KR2>  
 F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>  
 F:311-566/Domain: trypsin homology <TRY>  
 F:41-71-69-78-86-97-91-108-110-119-127-208-148-190-179-203-215-296-236-278-267-291-29  
 F:152-483/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental  
 F:310-311/Cleavage site: Arg-116 (plasmin, trypsin) #status experimental  
 F:357-408/Active site: His, Asp #status predicted  
 F:313/Active site: Ser #status experimental  
 Query Match 43.38; Score 216; DB 1; Length 562;  
 Best Local Similarity 47.08; Pred. NO. 3.7e-16;  
 Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;  
 QY 3 CYENGHFYRKASTDITMGRCPLWNSATVLDQGYHNRNSNALDGLGKKNCRPNDR 62  
 Db 215 CYFNGSAYRGTHTSLTESGASCLPWNMSMLIKGYTGAQNSAALGLGKKNCRPDGA 274  
 QY 63 RPYCYVGVGLKPLVQECNVHDC 85  
 Db 275 KPWCHVILKNRRLTWECYDVPSCS 297  
 RESULT 13  
 I38098  
 t-plasminogen activator precursor, inactive endothelial splice form - human  
 N:Alternate names: tissue plasminogen activator  
 C:Species: Homo sapiens (man)  
 C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 22-Jun-1999  
 C:Accession: I38098; S01678  
 R:Siebert, P.D.; Fong, K.  
 Nucleic Acids Res. 18, 1086, 1990  
 A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human en  
 A:Reference number: I38098; MUID:90192128; PMID:1969145  
 A:Accession: I38098  
 A:Status: translated from GB/EMBL/DBD



100



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 11:40:46 ; Search time 21 Seconds  
(without alignments)  
169.855 Million cell updates/sec

Title: PCT-US02-27855-1

Sequence: 1 KTCGEGNGHYRCKASNDTM.....YVGVGLKPLVGECHVHDCAD 86

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	494	99.0	431 1	UROK_HUMAN
2	459	92.0	433 1	UROK_PAPCY
3	409	82.0	442 1	UROK_PIG
4	400	80.2	432 1	UROK_RAT
5	383	76.8	433 1	UROK_BOVIN
6	381	76.4	433 1	UROK_MOUSE
7	236	47.3	431 1	UROB_DESRO
8	226	45.3	477 1	URR2_DESRO
9	226	45.3	559 1	TPA_RAT
10	218	43.7	394 1	URRG_DESRO
11	216	43.3	562 1	TPA_HUMAN
12	215	43.1	559 1	TPA_MOUSE
13	214	42.9	566 1	TPA_BOVIN
14	208	41.7	477 1	URR1_DESRO
15	203.5	40.8	434 1	UROK_CHICK
16	193	38.7	655 1	HGFL_HUMAN
17	188	37.7	653 1	HGFL_MOUSE
18	179.5	36.0	603 1	FAI2_CAVRO
19	170.5	34.2	615 1	FAI2_HUMAN
20	158	31.7	473 1	KREM_MOUSE
21	158	31.7	473 1	KREM_RAT
22	156	31.3	1420 1	APOA_MACMU
23	155.5	31.2	593 1	FAI2_BOVIN
24	155	31.1	475 1	KREM_HUMAN
25	155	31.1	4548 1	APOA_HUMAN
26	154	30.9	812 1	PLMN_BOVIN
27	153	30.7	790 1	PLMN_PIG
28	152.5	30.6	169 1	PLMN_RAT
29	149	29.9	810 1	PLMN_HUMAN
30	148.5	29.8	944 1	ROR2_MOUSE
31	147.5	29.6	810 1	PLMN_ERIEU
32	147	29.5	937 1	ROR1_HUMAN
33	147	29.5	937 1	ROR1_MOUSE

34	146	29.3	810 1	PLMN_MACMU	P12545 macaca mula
35	143.5	28.8	711 1	HGFL_HUMAN	P26927 homo sapien
36	143.5	28.8	943 1	ROR2_HUMAN	Q01974 homo sapien
37	143	28.7	333 1	PLMN_CANFA	P80009 canis famil
38	137	27.5	812 1	PLMN_MOUSE	P20918 mus musculu
39	133.5	26.8	728 1	HGF_MOUSE	Q08048 mus musculu
40	131.5	26.4	716 1	HGFL_MOUSE	P26928 mus musculu
41	131.5	26.4	728 1	HGF_HUMAN	P14210 homo sapien
42	131.5	26.4	728 1	HGF_RAT	P17945 rattus norv
43	126	25.3	325 1	PLMN_PETMA	P33574 petromyzon
44	125	25.1	622 1	THRB_HUMAN	P00734 homo sapien
45	124.5	24.9	618 1	THRB_MOUSE	P19221 mus musculu

## ALIGNMENTS

RESULT 1  
UROK\_HUMAN STANDARD: PRT; 431 AA.  
ID UROK\_HUMAN 015844: 016618: 0969W6;  
AC P00749: 015844: 016618: 0969W6;  
DT 21-JUN-1986 (Rel. 01, Created)  
DT 20-MAR-1987 (Rel. 04, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
DE (U-plasminogen activator).  
GN PLAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP MEDLINE=65215647; PubMed=2987867;  
RX Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;  
RT "The human urokinase-plasminogen activator gene and its promoter.";  
RL Nucleic Acids Res. 13:2759-2771(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,  
RA Steffens G.J., Heyneker H.L.;  
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";  
RL Biotechnology 3:923-929(1985).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=66056954; PubMed=2415429;  
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,  
RA Nishida M., Suyama T.;  
RT "Molecular cloning of cDNA coding for human preprourokinase.";  
RL Gene 36:183-188(1985).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=65203359; PubMed=3888571;  
RA Jacobs P., Cravador A., Lorian R., Brockly F., Colau B., Chuchana P.,  
RA van Elsen A., Herzog A., Bollen A.;  
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human preprourokinase cDNA.";  
RL DNA 4:139-146(1985).  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
RA Nickerson D.A.;  
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;  
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE OF 66-431 FROM N.A.  
RX MEDLINE=84272706; PubMed=6589620;  
RA Verde P., Stoppeilli M.P., Galeffi P., di Nocera P., Biasi F.;

RT "Identification and primary sequence of an unspliced human urokinase  
RT poly(A)+ RNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).  
[8]  
RP SEQUENCE OF 21-177.  
RX MEDLINE-83055084; PubMed-6754569;  
RA Gunler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,  
RA Flohe L.;  
RT "The primary structure of high molecular mass urokinase from human  
RT urine. The complete amino acid sequence of the A chain.";  
RL Hope-Seyler's Z. Physiol. Chem. 563:1155-1165(1982).  
[9]  
RP SEQUENCE OF 156-176 AND 179-224.  
RX MEDLINE-83003608; PubMed-6749491;  
RA Schaller J., Nick H., Rickli E.E., Gillissen D., Lergler W.,  
RA Studer R.O.;  
RT "Human low-molecular-weight urinary urokinase. Partial  
RT characterization and preliminary sequence data of the two polypeptide  
RT chains.";  
RL Eur. J. Biochem. 125:251-257(1982).  
[10]  
RP SEQUENCE OF 158-410.  
RX MEDLINE-83055099; PubMed-6754572;  
RA Steffens G.J., Gunler W.A., Otting F., Frankus E., Flohe L.;  
RT "The complete amino acid sequence of low molecular mass urokinase  
RT from human urine.";  
RL Hope-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).  
[11]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE-96000858; PubMed-8591045;  
RA Sprengon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,  
RA Dobson C.M., Stuart D.I., Jones E.Y.;  
RT "The crystal structure of the catalytic domain of human  
RT urokinase-type plasminogen activator.";  
RL Structure 3:681-691(1995).  
[12]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.  
RX MEDLINE-20266327; PubMed-10805774;  
RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,  
RA Bode W., Magdolen V., Huber R., Moroder L.;  
RT "(4-aminomethyl)phenylhydrazide derivatives as nonpeptidic highly  
RT selective inhibitors of human urokinase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).  
[13]  
RP STRUCTURE BY NMR.  
RX MEDLINE-86127526; PubMed-2536903;  
RA Oswald R.E., Boguski M.J., Bamberger M., Smith R.A.G., Dobson C.M.;  
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-  
RT dimensional NMR.";  
RL Nature 337:579-582(1989).  
[14]  
RP STRUCTURE BY NMR OF 67-155.  
RX MEDLINE-93003110; PubMed-1327118;  
RA Li X., Smith R.A.G., Dobson C.M.;  
RT "Sequential 1H NMR assignments and secondary structure of the kringle  
RT domain from urokinase.";  
RL Biochemistry 31:9562-9571(1992).  
[15]  
RP STRUCTURE BY NMR OF 67-155.  
RX MEDLINE-94149701; PubMed-8107091;  
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;  
RT "Solution structure of the kringle domain from urokinase-type  
RT plasminogen activator.";  
RL J. Mol. Biol. 235:1548-1559(1994).  
[16]  
RP VARIANT LEU-141.  
RX MEDLINE-96186279; PubMed-8652631;  
RA Yoshino M., Ushiyama T., Sakai M., Tamaki S., Hara H., Takahashi K.,  
RA Sawasaki Y., Hanada K.;  
RT "Characterization of single chain urokinase-type plasminogen  
RT activator with a novel amino-acid substitution in the kringle  
RT structure.";  
RL Biochim. Biophys. Acta 1293:83-89(1996).

RP [17]  
RP VARIANT LEU-141.  
RX MEDLINE-97218551; PubMed-9065988;  
RA Conne B., Berczy M., Bellin D.;  
RT "Detection of polymorphisms in the human urokinase-type plasminogen  
RT activator gene.";  
RL Thromb. Haemost. 77:434-435(1997).  
[18]  
RP ERRATUM.  
RA Conne B., Berczy M., Bellin D.;  
RL Thromb. Haemost. 78:973-973(1997).  
[19]  
RP VARIANT LEU-141.  
RX MEDLINE-97337920; PubMed-9194591;  
RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,  
RA Creutzburg S., Graeff H., Magdolen V.;  
RT "Mutational analysis of the genes encoding urokinase-type plasminogen  
RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";  
RL Electrophoresis 18:686-689(1997).  
CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR  
CC THERAPY OF THROMBOLYTIC DISORDERS.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in  
CC plasminogen to form plasmin.  
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS  
CC OF TWO CHAINS A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A  
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW  
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.  
CC -1- PHARMACEUTICAL: Available under the name AbboKinase (Abbott). Used  
CC in pulmonary embolism (see also fibrinolysis).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
CC  
DR EMBL: X02419; CAA26268.1; -  
DR EMBL: M15476; AAA61253.1; -  
DR EMBL: D00244; BAA00175.1; -  
DR EMBL: D11143; BAA01919.1; -  
DR EMBL: X02760; CAA26535.1; -  
DR EMBL: AF377330; AAK53822.1; -  
DR EMBL: BC013575; AAK13575.1; -  
DR EMBL: K03226; AAC97138.1; -  
DR EMBL: K02286; AAA61252.1; -  
DR EMBL: A21571; CAA01559.1; -  
DR EMBL: A18397; CAA01390.1; -  
DR PIR: A00931; UKHU  
DR PIR: A32974; A32974  
DR PDB: 1KDU; 31-OCT-93.  
DR PDB: 1IKW; 29-JAN-96.  
DR PDB: 1URK; 08-MAY-95.  
DR PDB: 1EYN; 17-MAY-00.  
DR MEROPS: S01.231; -  
DR GlycoSuiteDB: P00749; -  
DR Genew: HGNC:9052; PLAU.  
DR MIM: 191840; -  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR001254; Ser. protease-TRY.  
DR Pfam: PF00051; kringle; 1.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR ProDom: PD000395; kringle; 1.  
DR SMART: SM00181; EGF; 1.  
DR SMART: SM00130; KR; 1.  
DR SMART: SM00020; TRYP-SpC; 1.

Query Match 99.0%; Score 494; DB 1; Length 431;  
 Best Local Similarity 98.8%; Pred. No. 8,1e-50;  
 Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTCYBNGHFYRGKASTDTPMGRPCLPWNSATVLOOTYHAHRSNALOLGLGKHNCRPN 60  
 DB 68 KTCYBNGHFYRGKASTDTPMGRPCLPWNSATVLOOTYHAHRSNALOLGLGKHNCRPN 127  
 OY 61 RRRPMCYVGVGLKPLVQECMVHDCAD 86  
 DB 128 RRRPMCYVGVGLKPLVQECMVHDCAD 153

RESULT 2  
 UROK\_PAPCY STANDARD; PRT; 433 AA.  
 ID P16227;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)  
 GN PLAU.  
 OS Papio cynocephalus (Yellow baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Papio.  
 NC NCBL\_TaxID=9556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thoracic aorta;  
 RX MEDLINE=90287734; PubMed=2113276;  
 RA Au Y.P.T., Wang T.W., Clowes A.W.;  
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type  
 plasminogen activator.";  
 RT Nucleic Acids Res. 18:3411-3411(1990).  
 CC 1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 plasminogen to form plasmin.  
 CC 1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS  
 OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A  
 LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW  
 MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).  
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC 1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X51935; CAA36200.1; -  
 DR PIR: S14687; UKBAY.  
 DR HSSP: P00749; ILMW.  
 DR MEROPS: S01.231.-  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR001254; Ser.protease\_Try.  
 DR Pfam: PF00085; kringle\_1.  
 DR Pfam: PF00085; trypsin\_1.  
 DR PRINTS: PR00722; CHYMOTRYPsin.  
 DR PRODOM: PD000395; Kringle\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR SMART: SM00130; KR\_1.  
 DR SMART: SM00020; TRYp\_SPC\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; FALSE\_NTC.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.

DR PROSITE: PS50070; KRINGLE\_2; 1.  
 DR PROSITE: PS50240; TRYPsin\_DOM; 1.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 RM Kringle; EGF-like domain; Zymogen; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 433  
 FT CHAIN 21 176  
 FT CHAIN 155 176  
 FT CHAIN 178 433  
 FT CHAIN 26 62  
 FT DOMAIN 69 150  
 FT DOMAIN 151 177  
 FT DOMAIN 178 433  
 FT DISULFID 30 38  
 FT DISULFID 32 50  
 FT DISULFID 52 61  
 FT DISULFID 167 298  
 FT DISULFID 208 224  
 FT DISULFID 216 287  
 FT DISULFID 315 384  
 FT DISULFID 347 363  
 FT DISULFID 374 402  
 FT DISULFID 223 223  
 FT ACT\_SITE 274 274  
 FT ACT\_SITE 378 378  
 FT CARBOHD 324 324  
 SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CNC64;

Query Match 92.0%; Score 459; DB 1; Length 433;  
 Best Local Similarity 93.0%; Pred. No. 9.2e-46;  
 Matches 80; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KTCYBNGHFYRGKASTDTPMGRPCLPWNSATVLOOTYHAHRSNALOLGLGKHNCRPN 60  
 DB 67 KTCYBNGHFYRGKASTDTPMGRSCLAWSATVLOOTYHAHRSNALOLGLGKHNCRPN 126  
 OY 61 RRRPMCYVGVGLKPLVQECMVHDCAD 86  
 DB 127 RRRPMCYVGVGLKPLVQECMVHDCAD 152

RESULT 3  
 UROK\_PIG STANDARD; PRT; 442 AA.  
 ID P04185;  
 DT 20-MAR-1987 (Rel. 04, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)  
 GN (U-plasminogen activator).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 NC NCBL\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=85087954; PubMed=6096832;  
 RA Nagamine Y., Pearson D., Altus M.S., Reich E.;  
 RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";  
 RN Nucleic Acids Res. 12:9525-9541(1984).  
 RN [2]  
 RP REVISION TO 241.  
 RA Nagamine Y.;  
 RL Submitted (DEC-1986) to the PIR data bank.  
 CC 1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC 1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 CC 1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.



```

FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 289 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT CONFLICT 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 80.2%; Score 400; DB 1; Length 432;
Best Local Similarity 80.0%; Pred. No. 6.2e-39;
Matches 68; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDWMGRPCLPWNSATVYLQOTYHAHRSNALQGLGKHNCRNDN 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 68 KTCYHNGSGSYRGKANTDGRPCLPWNSATVYLQOTYHAHRSNALQGLGKHNCRNDN 127
QY 61 RRRPWCYVGVGLKPLVQECMVHDC 85
    :|||||:|||||:|||||:
DB 128 QRRPWCYVGLKQVQECMVHDCS 152

RESULT 5
UROC_BOVIN STANDARD; PRT; 433 AA.
AC 005589; 028209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Aortic endothelium;
RA MEDLINE=93216119; PubMed=8385052;
RA Kraetzschmar J, Haendler B, Kojima S, Rifkin D.B.,
RA Schlaunung W.-D.,
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE-Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

```

```

CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03546; AA551419.1; -
DR EMBL; X85801; CA59796.1; -
DR PIR; JN0560; JN0560.
DR HSSP; P00749; 1LMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 179 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 181 433 CHAIN A (BY SIMILARITY).
FT DOMAIN 29 65 CHAIN B (BY SIMILARITY).
FT DOMAIN 72 153 EGF-LIKE.
FT DOMAIN 154 180 KRINGLE.
FT DISULFID 181 433 CONNECTING PEPTIDE.
FT DISULFID 33 41 SERINE PROTEASE.
FT DISULFID 35 64 BY SIMILARITY.
FT DISULFID 55 63 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 76.8%; Score 383; DB 1; Length 433;
Best Local Similarity 75.3%; Pred. No. 5.8e-37;
Matches 64; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDWMGRPCLPWNSATVYLQOTYHAHRSNALQGLGKHNCRNDN 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 70 KTCYQNGSHSYRGKANTDGRPCLPWNSATVYLQOTYHAHRSNALQGLGKHNCRNDN 129
QY 61 RRRPWCYVGVGLKPLVQECMVHDC 85
    :|||||:|||||:|||||:
DB 130 QRRPWCYVGLKQVQECMVHDCS 154

RESULT 6
UROC_MOUSE STANDARD; PRT; 433 AA.
AC P06869;

```

DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)  
 DE (U-plasminogen activator).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85179474; PubMed=2985383;  
 RA Belin D., Vassalli J.-D., Combejean C., Godeau F., Nagamine Y.,  
 RA Reich E., Kocher H.P., Duvoisin R.M.;  
 RT Cloning, nucleotide sequencing and expression of cDNAs encoding  
 RT mouse urokinase-type plasminogen activator.\*;  
 RL Eur. J. Biochem. 148:225-232(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88163489; PubMed=2831940;  
 RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;  
 RT "The murine urokinase-type plasminogen activator gene.\*";  
 RL Biochemistry 26:8270-8278(1987).  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS  
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A  
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
 CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
 CC -----  
 CC EMBL: X02389; CAA26231.1;  
 CC EMBL: M17922; AAA0539.1;  
 CC PIR: A24615; UKMS.  
 CC PIR: A29420; A29420.  
 CC HSSP: P00749; 1KDU.  
 CC DR HSSP: S01\_231;  
 CC DR MGD: MGT:97611; Plau.  
 CC DR InterPro: IPR001314; Chymotrypsin.  
 CC DR InterPro: IPR000361; EGF-like.  
 CC DR InterPro: IPR000001; Kringle.  
 CC DR InterPro: IPR001254; Ser-protease\_Try.  
 CC DR Pfam: PF00051; Kringle; 1.  
 CC DR Pfam: PF00089; Trypsin; 1.  
 CC DR PRINTS: PR00722; CHYMOTRYPSIN.  
 CC DR ProDom: PD000395; Kringle; 1.  
 CC DR SMART: SM00181; EGF; 1.  
 CC DR SMART: SM00130; Kr; 1.  
 CC DR SMART: SM00020; TRP-Spec; 1.  
 CC DR PROSITE: PS00022; EGF\_1; 1.  
 CC DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 CC DR PROSITE: PS00021; KRINGLE\_1; 1.  
 CC DR PROSITE: PS00070; KRINGLE\_2; 1.  
 CC DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 CC DR PROSITE: PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
 CC DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 CC KM plasminogen activator; Hydrolase; Serine protease; Glycoprotein;  
 KM Kringle; EGF-like domain; Zymogen; Signal.  
 FM SIGNAL  
 FT CHAIN 1 20  
 FT CHAIN 21 433  
 FT CHAIN 157 178  
 FT CHAIN 157 178  
 CHAIN 157 178  
 SHORT A CHAIN (A1).

FT CHAIN 180 433 CHAIN B (BY SIMILARITY).  
 FT DOMAIN 28 64 EGF-LIKE.  
 FT DOMAIN 71 152 KRINGLE.  
 FT DOMAIN 153 179 CONNECTING PEPTIDE.  
 FT DOMAIN 180 433 SERINE PROTEASE.  
 FT DISULFID 32 40 BY SIMILARITY.  
 FT DISULFID 34 52 BY SIMILARITY.  
 FT DISULFID 54 63 BY SIMILARITY.  
 FT DISULFID 169 199 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 211 227 BY SIMILARITY.  
 FT DISULFID 219 290 BY SIMILARITY.  
 FT DISULFID 315 384 BY SIMILARITY.  
 FT DISULFID 347 363 BY SIMILARITY.  
 FT DISULFID 374 402 BY SIMILARITY.  
 FT ACT SITE 226 226 CHARGE RELAY SYSTEM.  
 FT ACT SITE 277 277 CHARGE RELAY SYSTEM.  
 FT ACT SITE 378 378 CHARGE RELAY SYSTEM.  
 SQ SEQUENCE 433 AA; 48268 MW; A99C35F625043F9 CRC64;  
 Query Match 76.4%; Score 381; DB 1; Length 433;  
 Best Local Similarity 72.9%; Pred. No. 9.9e-37; Indels 0; Gaps 0;  
 Matches 62; Conservative 11; Mismatches 12;  
 QY 1 KTCYEGNGHFRKASTDTPGRPCLEPWSATVLOQTYSRHSNALQGLGRHNYCRNPDN 60  
 DB 69 KTCYHNGSDSYRKANTDTKGRPCLEPWSATVLOQTYSRHSNALQGLGRHNYCRNPDN 128  
 QY 61 RRRPWCYVQVGLKPLVQECWHDCA 85  
 DB 129 QRRPWCYVQVGLKPLVQECWHDCA 153  
 RESULT 7  
 URTB\_DESRO STANDARD; PRT; 431 AA.  
 AC P98121;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA  
 DE beta).  
 OS Desmodus rotundus (Vampire bat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
 OC Desmodontinae; Desmodus.  
 OX NCBI\_TaxID=9430;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Salivary gland;  
 RX MEDLINE=92039036; PubMed=1937019;  
 RA Kretschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
 RA Alagon A., Donner P., Schleuning W.D.;  
 RT "The plasminogen activator family from the salivary gland of the  
 RT vampire bat Desmodus rotundus: cloning and expression.\*";  
 RL Gene 105:229-237(1991).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=9339059; PubMed=1309059;  
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
 RA Kretschmar J., Haendler B., Langer G., Baldus B., Wilt W.,  
 RA Donner P.;  
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
 RT vampire bat): unique fibrin specificity.\*";  
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS  
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC  
 CC AGENT.  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: M63986; AAA31594.1; -  
CC HSSP: P98119; 1A51.  
CC MEROPS: S01.239; -  
CC InterPro: IPR001314; Chymotrypsin.  
CC InterPro: IPR000561; EGF-like.  
CC InterPro: IPR000001; Kringle.  
CC InterPro: IPR001254; Ser-protease\_Try.  
CC Pfam: PF00089; EGF\_1.  
CC Pfam: PF00051; kringle\_1.  
CC Pfam: PF00089; trypsin\_1.  
CC PRINTS: PR00722; CHYMOTRYPSIN.  
CC PRINTS: PR00018; KRINGLE.  
CC ProDom: PD000395; Kringle; 1.  
CC SMART: SM00181; EGF; 1.  
CC SMART: SM00130; KR; 1.  
CC SMART: SM00020; TRYP-SPC; 1.  
CC PROSITE: PS00022; EGF\_1; 1.  
CC PROSITE: PS01186; EGF\_2; 1.  
CC PROSITE: PS00021; KRINGLE\_1; 1.  
CC PROSITE: PS00021; KRINGLE\_2; 1.  
CC PROSITE: PS50240; TRYPsin\_DOM; 1.  
CC PROSITE: PS00134; TRYPsin\_HIS; 1.  
CC PROSITE: PS00135; TRYPsin\_SER; 1.  
CC Plasminogen activator: Hydrolyase; Serine protease; Glycoprotein;  
CC Kringle; EGF-like domain; Signal; Multigene family.  
CC SIGNAL 1 36  
CC CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.  
CC DOMAIN 37 75 EGF-LIKE.  
CC DOMAIN 82 163 KRINGLE.  
CC ACT\_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).  
CC ACT\_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).  
CC ACT\_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).  
CC DISULFID 46 52 BY SIMILARITY.  
CC DISULFID 46 63 BY SIMILARITY.  
CC DISULFID 65 74 BY SIMILARITY.  
CC DISULFID 82 163 BY SIMILARITY.  
CC DISULFID 103 145 BY SIMILARITY.  
CC DISULFID 134 158 BY SIMILARITY.  
CC DISULFID 168 299 BY SIMILARITY.  
CC DISULFID 211 227 BY SIMILARITY.  
CC DISULFID 219 288 BY SIMILARITY.  
CC DISULFID 313 388 BY SIMILARITY.  
CC DISULFID 345 361 BY SIMILARITY.  
CC DISULFID 378 406 BY SIMILARITY.  
CC CARBOHYD 139 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;  
-----  
Query Match 47.3%; Score 236; DB 1; Length 431;  
Best Local Similarity 48.8%; Pred. No. 6, 1e-20;  
Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;  
-----  
QY 2 TCEGSHHYRGKASTDTMGRCPLPNVSAATVLOQTYHAHRNANLQGLGKHNVCRRPDKR 61  
DB 81 TCKKDGCVYTRGTWSTSEGAQCTINNSMLTRITNGRSADLTTLGLNHNICRPDNN 140  
QY 62 RRPWCYVQGLKPLVQECWHDCA 85  
DB 141 SKPMCVIVAKSKFLEFCVPCS 164  
-----  
RESULT 8  
URT2\_DESRO

-----  
ID URT2\_DESRO STANDARD; PRT; 477 AA.  
AC P15638;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA  
DE alpha-2) (BAM-PA) (T-plasminogen activator).  
OS Desmodus rotundus (Vampire bat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
OC Desmodontinae; Desmodus.  
OX NCBI\_TaxID=9430;  
RN [1]  
RP TISSUE=Salivary gland;  
RC MEDLINE=92039036; PubMed=1937019;  
RA Krietzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
RA Alagon A., Donner P., Schleuning W.D.;  
RT "The plasminogen activator family from the salivary gland of the  
RT vampire bat Desmodus rotundus: cloning and expression.";  
RL Gene 105:229-237(1991).  
[2]  
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RP TISSUE=Salivary gland;  
RC MEDLINE=90036867; PubMed=2509450;  
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,  
RA Regester R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;  
RT "Isolation, characterization, and cDNA cloning of a vampire bat  
RT salivary plasminogen activator.";  
RL J. Biol. Chem. 264:17947-17952(1989).  
[3]  
RN CHARACTERIZATION.  
RP MEDLINE=93393059; PubMed=1309059;  
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
RA Krietzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
RA Donner P.;  
RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
RT vampire bat): unique fibrin specificity.";  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992)  
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS  
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC  
CC AGENT.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN  
CC THE PRESENCE OF FIBRIN I.  
CC -1- SUBUNIT: MONOMER.  
CC -1- DOMAIN: THE FIBRONECTIN TYPE-1 DOMAIN MEDIATES BINDING TO FIBRIN,  
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED  
CC STIMULATION OF ACTIVITY.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: M63986; AAA31593.1; -  
CC EMBL: J05082; AAA31596.1; -  
CC PIR: A34369; A34369.  
CC HSSP: P98119; 1A51.  
CC MEROPS: S01.232; -  
CC InterPro: IPR001314; Chymotrypsin.  
CC InterPro: IPR000561; EGF-like.  
CC InterPro: IPR000083; Fibrinctn.  
CC InterPro: IPR000001; Kringle.  
CC InterPro: IPR001254; Ser-protease\_Try.  
-----

```

DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00039; fnl; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00186; KRINGLE.
DR Pfam: PF000395; kringle; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00058; FNL; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; tryp_spc; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KR kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT ACQ_SITE 272 272
FT ACQ_SITE 321 321
FT ACQ_SITE 428 428
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 109
FT DISULFID 92 120
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 345
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 185 185
FT CARBOHYD 398 398
FT CONFLICT 403 403
FT CONFLICT 417 417
FT CONFLICT 435 435
SO SEQUENCE 477 AA; 53719 MW; 17486555C05077C CRC64;

Query Match 47.38; Score 236; DB 1; Length 477;
Best Local Similarity 48.86; Pred. No. 6.7e-20;
Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 2 TCYEGNGHFYKSGASTDTMGPRCLPMNSATVLOQTYHAHNSNALQGLGRHNTCRNPNDNR 61
Db 127 TCYDQGTGTYGTWSTESGACINMNSNLTTRTYNGRRSDAITLGLGHNNYCRNDNN 186
OY 62 RRPWCYVOVGLKPLVQECMVHDC 85
Db 187 SKPMCYVIAKSKFILEFCSPVCS 210

```

```

GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
(1)
RP SEQUENCE FROM N.A. PubMed-3148445;
RX MEDLINE-89170114;
RA Ny T., Leonardsson G., Hsueh A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator."
RL DNA 7:671-677(1988).
(2)
RP SEQUENCE FROM N.A.
RX MEDLINE-90130448; PubMed-2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the rRNA-less rat tissue-type plasminogen activator
RT gene. Species-specific sequence divergences in the promoter predict
RT differences in regulation of gene expression."
RL J. Biol. Chem. 265:2022-2027(1990).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOSIN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: M23697; AAA41812.1; -
DR EMBL: M31197; AAA42261.1; JOINED.
DR EMBL: M31185; AAA42261.1; JOINED.
DR EMBL: M31186; AAA42261.1; JOINED.
DR EMBL: M31187; AAA42261.1; JOINED.
DR EMBL: M31188; AAA42261.1; JOINED.
DR EMBL: M31189; AAA42261.1; JOINED.
DR EMBL: M31190; AAA42261.1; JOINED.
DR EMBL: M31191; AAA42261.1; JOINED.
DR EMBL: M31192; AAA42261.1; JOINED.
DR EMBL: M31193; AAA42261.1; JOINED.
DR EMBL: M31194; AAA42261.1; JOINED.
DR EMBL: M31195; AAA42261.1; JOINED.
DR EMBL: M31196; AAA42261.1; JOINED.
DR EMBL: A19618; CA001482.1; -
DR PIR: A31597; A31597.
DR PIR: A35029; A35029.
DR HSR: P00750; 1RTP.
DR MEROPS: S01.232; -
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000083; Fibrinctn1.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 1.

```



```

DR Pfam: PF00039; fnl; 1.
DR Pfam: PF00051; kringle; 2.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR Prodom: PD000395; Kringle; 2.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00058; FNI; 1.
DR SMART: SM00130; KR; 2.
DR SMART: SM00020; TRYP-Spec; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS0186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS00240; TRYP-SIN_DOM; 1.
DR PROSITE: PS00134; TRYP-SIN_HIS; 1.
DR PROSITE: PS00135; TRYP-SIN_SER; 1.
KW Plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
KV Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT DOMAIN 36 78 FIBRONECTIN TYPE-1.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 KRINGLE 1.
FT DOMAIN 213 294 KRINGLE 2.
FT DOMAIN 309 559 SERINE PROTEASE.
FT ACT_SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 68 68 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 480 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 380 380 E -> K (IN REF. 1).
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;
Query Match 45.3%; Score 226; DB 1; Length 559;
Best Local Similarity 49.4%; Pred. No. 1, 1e-18;
Matches 41; Conservative 9; Mismatches 33; Indels 0; Gaps 0;

```

```

AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSFA
DE gamma).
OC Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kretschmar J., Haendler B., Langer G., Boisdol W., Bringmann P.,
RA Alagon A., Donner P., Schlenning W.D., the salivary gland of the
RT "The plasminogen activator family from the salivary gland of the
RL vampire bat Desmodus rotundus: cloning and expression."
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schlenning W.-D., Alagon A., Boisdol W., Bringmann P., Petri T.,
RA Kretschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity."
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
-1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
AGENT.
-1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
-1- SUBUNIT: MONOMER.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib.ch).
-----
DR EMBL: M63990; AAA31595.1; -.
DR HSSP: P98119; IAS1.
DR MEROPS: S01.239; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam: PF00051; kringle; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR Prodom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYP-Spec; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYP-SIN_DOM; 1.
DR PROSITE: PS00134; TRYP-SIN_HIS; 1.
DR PROSITE: PS00135; TRYP-SIN_SER; 1.
KW Plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
KV Kringle; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT DOMAIN 142 394 KRINGLE.
FT ACT_SITE 189 189 SERINE PROTEASE.
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).

```

FT DISULFID 45 126 BY SIMILARITY.  
 FT DISULFID 66 108 BY SIMILARITY.  
 FT DISULFID 97 121 BY SIMILARITY.  
 FT DISULFID 131 262 BY SIMILARITY.  
 FT DISULFID 174 190 BY SIMILARITY.  
 FT DISULFID 182 251 BY SIMILARITY.  
 FT DISULFID 276 351 BY SIMILARITY.  
 FT DISULFID 308 324 BY SIMILARITY.  
 FT DISULFID 341 369 BY SIMILARITY.  
 FT CARBOHYD 315 315 N-LINKED (GLCNAC...)  
 SO SEQUENCE 394 AA: 44105 MM: 9CCD6F52F3D81FCD CR64;

Query Match 43.7%; Score 218; DB 1; Length 394;  
 Best Local Similarity 45.2%; Pred. No. 6.7e-18;  
 Matches 38; Conservative 12; Mismatches 34; Indels 0; Gaps 0;

QY 2 TCYGNHGFTRKASTDTMGRCPLPNWATVLOQTYHAHRSNALDGLGKHNCRNDNR 61  
 DB 44 TCYKDGVTYRGWSTSESGAQCINMNSNLLIRRYNGRMPKAVKLGHNHYCRNDPGA 103  
 QY 62 RRPWCYVQVGLKPLVQECMVHDC 85  
 DB 104 SKPCYVYKARKFTSESCVPCVS 127

RESULT 11  
 ID TPA\_HUMAN STANDARD; PRT; 562 AA.  
 AC P00750; Q15103;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last annotation update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
 GN (t-PA) (t-plasminogen activator) (Alteplase) (Keteplase).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Melanoma;  
 RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,  
 RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,  
 RA Goeddel D.V., Collen D.;  
 RT "Cloning and expression of human tissue-type plasminogen activator  
 RT cDNA in E. coli.";  
 RL Nature 301:214-221(1983).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Fetal Lung;  
 RA MEDLINE-83115262; PubMed-6337343;  
 RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niva M.;  
 RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA  
 RT from human fetal lung cells.";  
 RL Nucleic Acids Res. 16:5695-5695(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-88054470; PubMed-2824147;  
 RA Reddy V.B., Garramone A.J., Sasaki H., Wei C.-M., Watkins P., Galli J.,  
 RA Hsiung N.;  
 RT "Expression of human uterine tissue-type plasminogen activator in  
 RT mouse cells using BPV vectors.";  
 RL DNA 6:461-472(1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-86196143; PubMed-3009482;  
 RA Friezner Degen S.J., Rajput B., Reich E.;  
 RT "The human tissue plasminogen activator gene.";  
 RL J. Biol. Chem. 261:6972-6985(1986).  
 RN [5]  
 RP SEQUENCE FROM N.A.

RX MEDLINE-84298137; PubMed-6089198;  
 RA "The structure of the human tissue-type plasminogen activator gene:  
 RA correlation of intron and exon structures to functional and  
 RA structural domains.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-86284200; PubMed-3090401;  
 RA Harris T.J., Patel T., Marston F.A., Little S., Entage J.S.,  
 RA Odenacker G., Volckert G., Rombaux W., Billiau A., Somer P.;  
 RT "Cloning of cDNA coding for human tissue-type plasminogen activator  
 RT and its expression in Escherichia coli.";  
 RL Mol. Biol. Med. 3:279-292(1986).  
 RN [7]  
 RP SEQUENCE OF 212-361 FROM N.A.  
 RC MEDLINE-83169656; PubMed-6572897;  
 RA Edlund T., Ny T., Raandy M., Heden L.-O., Palm G., Holmgren E.,  
 RA Josephson S.;  
 RT "Isolation of cDNA sequences coding for a part of human tissue  
 RT plasminogen activator.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).  
 RN [8]  
 RP SEQUENCE OF 1-36 FROM N.A.  
 RC MEDLINE-85289338; PubMed-3161893;  
 RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,  
 RA Schlenker W.-D.;  
 RT "Isolation and characterization of the human tissue-type plasminogen  
 RT activator structural gene including its 5' flanking region.";  
 RL J. Biol. Chem. 260:11223-11230(1985).  
 RN [9]  
 RP SEQUENCE OF 31-562 FROM N.A.  
 RC MEDLINE-91291340; PubMed-1368681;  
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;  
 RT "Purification and characterization of tissue plasminogen activator  
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";  
 RL Agric. Biol. Chem. 55:1225-1232(1991).  
 RN [10]  
 RP SEQUENCE OF 36-562.  
 RC TISSUE-Melanoma;  
 RA MEDLINE-85000468; PubMed-6433976;  
 RA Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;  
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly  
 RT derived amino acid sequence, identify the active site serine residue,  
 RT establish glycosylation sites, and localize variant differences.";  
 RL Biochemistry 23:3701-3707(1984).  
 RN [11]  
 RP SEQUENCE OF 33-52 AND 311-330.  
 RC TISSUE-Melanoma;  
 RA MEDLINE-83209620; PubMed-6682760;  
 RA Wallen P., Pohl G., Bergsdorf N., Raandy M., Ny T., Joernvall H.;  
 RT "Purification and characterization of a melanoma cell plasminogen  
 RT activator.";  
 RL Eur. J. Biochem. 132:681-686(1983).  
 RN [12]  
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
 RC TISSUE-Umbilical vein;  
 RA MEDLINE-90192129; PubMed-2107528;  
 RA Siebert P.D., Fong K.;  
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from  
 RT human endothelial cells.";  
 RL Nucleic Acids Res. 18:1086-1086(1990).  
 RN [13]  
 RP STRUCTURE OF CARBOHYDRATES.  
 RC MEDLINE-90092112; PubMed-2511186;  
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;  
 RT "Carbohydrate structure of recombinant human uterine tissue  
 RT plasminogen activator expressed in mouse epithelial cells.";  
 RL Eur. J. Biochem. 186:273-286(1989).  
 RN [14]  
 RP CARBOHYDRATE-LINKAGE SITE THR-96.  
 RC MEDLINE-91159408; PubMed-1900431;  
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;

RT "Tissue plasminogen activator has an O-linked fucose attached to  
 RT theone-61 in the epidermal growth factor domain."  
 RL Biochemistry 30:2311-2314(1991).  
 RN [15]  
 RP DISULFIDE BONDS IN KRINGLE 2.  
 RX MEDLINE-91244765; PubMed-1645336;  
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;  
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue  
 RT plasminogen activator produced in *Escherichia coli*."  
 RL J. Biol. Chem. 266:10070-10072(1991).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.  
 RX MEDLINE-96200985; PubMed-8613982;  
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,  
 RT "The 2.3 A crystal structure of the catalytic domain of recombinant  
 RT two-chain human tissue-type plasminogen activator."  
 RL J. Mol. Biol. 258:117-135(1996).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.  
 RX MEDLINE-97449126; PubMed-9305622;  
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,  
 RT "Lysine 156 promotes the anomalous proenzyme activity of tpa: X-ray  
 RT crystal structure of single-chain human tpa."  
 RL EMBO J. 16:4797-4805(1997).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.  
 RX MEDLINE-92118803; PubMed-1310033;  
 RA de Vos A., Ulsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,  
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen  
 RT activator at 2.4-A resolution."  
 RL Biochemistry 31:270-279(1992).  
 RN [19]  
 RP STRUCTURE BY NMR OF KRINGLE 2.  
 RX MEDLINE-90122799; PubMed-2558718;  
 RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
 RT "1H NMR structural characterization of a recombinant kringle 2 domain  
 RT from human tissue-type plasminogen activator."  
 RL Biochemistry 28:9350-9360(1989).  
 RN [20]  
 RP STRUCTURE BY NMR OF KRINGLE 2.  
 RX MEDLINE-91200042; PubMed-1901789;  
 RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR  
 RT assignments and secondary structure."  
 RL Eur. J. Biochem. 197:155-165(1991).  
 RN [21]  
 RP STRUCTURE BY NMR OF KRINGLE 2.  
 RX MEDLINE-92106329; PubMed-1762144;  
 RA Byeon I.-J.L., Llinas M.;  
 RT "Solution structure of the tissue-type plasminogen activator kringle  
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic  
 RT drug."  
 RL J. Mol. Biol. 222:1035-1051(1991).  
 RN [22]  
 RP STRUCTURE BY NMR OF 38-85.  
 RX MEDLINE-92292163; PubMed-1602484;  
 RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,  
 RT "Solution structure of the fibrin binding finger domain of  
 RT tissue-type plasminogen activator determined by 1H nuclear magnetic  
 RT resonance."  
 RL J. Mol. Biol. 225:821-833(1992).  
 RN [23]  
 RP STRUCTURE BY NMR OF 36-126.  
 RX MEDLINE-96027104; PubMed-7582899;  
 RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;  
 RT "The solution structure and backbone dynamics of the fibronectin type  
 RT I and epidermal growth factor-like pair of modules of tissue-type  
 RT plasminogen activator."  
 RL Structure 3:823-833(1995).

CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN  
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY  
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT  
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND  
 CC MANY OTHER PHYSIOLOGICAL EVENTS.  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in  
 CC plasminogen to form plasmin.  
 CC -1- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION  
 CC  
 CC Query Match 43.38; Score 216; DB 1; Length 562;  
 CC Best Local Similarity 47.08; Pred. No. 1.6e-17;  
 CC Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;  
 CC  
 CC 3 CYEGNGHFYRGKASTDFMGRCLEPMNSATVYDQTYHARSNALDLGKHNCRPNRR 62  
 CC 215 CYFNGSAYRGTHSLTEGASCLPWNMSILLGKTYTQNSAQLGKHNCRPNR 274  
 CC  
 CC 63 RPWCYVOYGLKPLVQECMVHDC 85  
 CC 275 KPWCHVLKRNRLTWECVDPSCS 297  
 CC  
 CC RESULT 12  
 CC TPA\_MOUSE STANDARD; PRT; 559 AA.  
 CC ID TPA\_MOUSE  
 CC AC P11214;  
 CC DT 01-JUL-1989 (Rel. 11, Created)  
 CC DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tpa)  
 CC DE (t-PA) (t-plasminogen activator).  
 CC GN PLAT.  
 CC OS Mus musculus (Mouse).  
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC OX NCBI\_TaxID=10090;  
 CC RX MEDLINE-88087303; PubMed-2826484;  
 CC RA Rickles R.J., Darrow A.L., Strickland S.;  
 RT "Molecular cloning of complementary DNA to mouse tissue plasminogen  
 RT activator mRNA and its expression during F9 teratocarcinoma cell  
 RT differentiation."  
 CC CC J. Biol. Chem. 263:1563-1569(1988).  
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN  
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY  
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT  
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND  
 CC MANY OTHER PHYSIOLOGICAL EVENTS.  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in  
 CC plasminogen to form plasmin.  
 CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE  
 CC BOND.  
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.  
 CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
 CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
 CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A  
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: J03520; AAA40470.1; -.

PTIR	A29941	A29941	
DR	HSSP	P00750	1A5H.
DR	MEROPS	S01232	.
DR	MCD	NCI_97610	P1t.
DR	InterPro	IPR00314	Glycocytpsins.
DR	InterPro	IPR000561	EGF-like.
DR	InterPro	IPR000083	Fibronctnl.
DR	InterPro	IPR000001	Ser_protease_Try.
DR	Pfam	PF00008	EGF_1.
DR	Pfam	PF00039	fn1_1.
DR	Pfam	PF00051	kringle_1.
DR	Pfam	PF00089	trypsin_1.
DR	PRINTS	PR00722	CHYOTRYPSIN.
DR	PRINTS	P00018	KRINGLE.
DR	ProDom	P000395	Kringle_2.
DR	SMART	SM00181	EGF_1.
DR	SMART	SM00058	FN1_1.
DR	SMART	SM00130	Kr_2.
DR	SMART	SM00020	Tryp_Spc_1.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.
DR	PROSITE	PS00070	KRINGLE_1_2.
DR	PROSITE	PS00020	KRINGLE_2_2.
DR	PROSITE	PS02040	TRYPSIN_DOM_1.
DR	PROSITE	PS00134	TRYPSIN_HIS_1.
DR	PROSITE	PS00135	TRYPSIN_SER_1.
DR	Plasma	kringle	EGF-like domain; Repeat; Signal.
DR	PROSITE	PS00022	EGF_1_1.
DR	PROSITE	PS01186	EGF_2_1.
DR	PROSITE	PS01253	FIBRONECTIN_1_1.

Query Match: 43.1%; Score 215; DB 1; Length 559.  
Best Local Similarity: 47.0%; Pred. No. 2,16-17;  
Matches 39; Conservative 10; Mismatches 34; Indels 0; Gaps 0.

2 TCIEGNGHPRFGASTDIMGKPCLPWNSATYLOQTYHAHRNALOLGLGHNYCRANDNR 61  
|||||  
Db 123 TCEEGDITRTGWSTAESGAECIMNNSVSLKKYNNARNNAIKLIGHNACRNDRD 182  
|||||

Qy	62	RBPACVYGVGLKPLVOCBWHDC	84
		:	
Dd	183	LKPBCIVFKAGKTTEFCSTPAC	205

RESULT	ID	TPA_BOVIN	STANDARD:	PRT:	566 AA.
	DT	01-NOV-1997	(Rel. 35, Created)		
	DT	01-NOV-1997	(Rel. 35, Last sequence update)		
	DE	15-JUN-2002	(Rel. 41, Last annotation update)		
	DE	(t-pa)	(t-plasminogen activator precursor (EC 3.4.21.68) (tPA))		
	GN	PLAT.			
	OS	Bos taurus (Bovins).			
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
	OX	Bovidae; Bovinae; Bos.			
	RM	NCL_TaxID=9913,			
	RP	SEQUENCE FROM N.A.			
	RC	TISSUE=Kidney;			
	RA	Kavn P., Berglund L., Petersen T.E.;			
	RT	"Cloning and characterization of the bovine plasminogen activators uPA			
	RL	and tPA.";			
	RL	Int. Dairy J. 5:605-617(1995).			
	-I-	FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, XENOGEN PLASMINOGEN			
	CC	TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY			
	CC	CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT			
	CC	ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND			
	CC	MANY OTHER PHYSIOPATHOLOGICAL EVENTS.			
	-I-	CATALYTIC ACTIVITY: Specific cleavage of Arg-I-Val bond in			
	CC	plasminogen to form plasmin.			
	-I-	SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE			
	CC	BOND.			
	-I-	SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.			
	CC	-I- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER			
	CC	PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER			
	CC	ARG-31 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XI.			
	-I-	MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A			
	CC	CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.			
	-I-	SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
	CC	-I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
	CC	-I- SIMILARITY: CONTAINS 1 FIBRINOCTIN TYPE I DOMAIN.			
	CC	-I- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.			
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
	CC	the European Bioinformatics Institute. There are no restrictions on its			
	CC	use by non-profit institutions as long as its content is in no way			
	CC	modified and this statement is not removed. Usage by and for commercial			
	CC	entities requires a license agreement (see <a href="http://www.isb-slb.ch/announcement/">http://www.isb-slb.ch/announcement/</a>			
	CC	or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).			
	DR	EMBL: X85800; CAAS9795.1; .			
	DR	HSSP: P00750; IRPF.			
	DR	MEROPS: S01_232; .			
	DR	InterPro: IPRO01314; Chymotrypsin.			
	DR	InterPro: IPRO00561; EGF-like.			
	DR	InterPro: IPRO00083; Fibrinctn1.			
	DR	InterPro: IPRO00001; Kringle.			
	DR	InterPro: IPRO01254; Ser_protease-Try.			
	DR	Pfam: PF00008; EGF; 1.			
	DR	Pfam: PF00039; fnl; 1.			
	DR	Pfam: PF00051; kringle; 2.			
	DR	Pfam: PF00089; trypsin; 1.			
	DR	PRINTS: PR00722; CHYMOTRYPSIN.			
	DR	PRINTS: PR00018; KRINGLE.			
	DR	ProDom: PD000395; kringle; 2.			
	DR	SMART: SM00181; EGF; 1.			
	DR	SMART: SM00058; FN1; 1.			

DR	SMART: SMO0130; KR: 2.
DR	SMART: SMO0020; TYP_SPC: 1.
DR	PROSITE; PS00022; BGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.
DR	PROSITE; PS00021; KRINGLE_1; 1.
DR	PROSITE; PS00070; KRINGLE_2; 2.
DR	PROSITE; PSS0240; TRYPsin_DOM. 1.
DR	PROSITE; PSS0134; TRYPsin_HIS; 1.
DR	PROSITE; PS00135; TRYPsin_SER; 1.
KW	Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KM	Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT	SIGNAL
FT	PROPEP
FT	CHAIN
FT	CHAIN
FT	CHAIN
FT	CHAIN
FT	CHAIN
FT	CHAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	ACT_SITE
FT	ACT_SITE
FT	ACT_SITE
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	CARBOHYD
FT	CARBOHYD
SO	SEQUENCE
Query Match	Best Local Similarity
Matches	Conservative
OY	2 TCYEENGHFYRGKASTDTMGRCPLVNSATVLOQGYHHRNSALDLGLKHNYCRNPDR 61
Db	127 TCYNKGVAIRNRWTMTAEGACCAWMNMSGLAMRKYSGRPRNAIRLGLGNHNHCANPDG 186
OY	62 RRPWCYVOGLKPLVOECMHDC A 85
Db	187 SKPMCYVEFKAGKYISEFCSTPACA 210
RESULT 14	
ID	UPTI_DESRO
ID	P98119.
DT	01-FEB-1996 (Rel. 33, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA alpha-1).
OS	Dermodys rotundus (Vampire bat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;

Desmodontinae: Desmodus.  
NCBI\_TaxID=9430;  
[1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Salivary gland;  
RX MEDLINE=92039036; PubMed=1937019;  
RA Kretschnmar J., Haendler B., Langer G., Boldol W., Bringmann P.,  
Alagon A., Donner P., Schlenning W.D.;  
RT "The plasminogen activator family from the salivary gland of the  
RT vampire bat Desmodus rotundus: cloning and expression."  
RL Gene 105:229-237(1991).  
[2]  
RP CHARACTERIZATION  
RX MEDLINE=93393059; PubMed=1309059;  
RA Schlenning W.-D., Alagon A., Boldol W., Bringmann P., Petri T.,  
RA Kretschnmar J., Haendler B., Langer G., Baldus B., Wilt W.,  
RA Donner P.;  
RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
RT vampire bat): unique fibrin specificity";  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
[3]  
RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RP TISSUE=Salivary gland;  
RX MEDLINE=98022741; PubMed=9354616;  
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,  
RA Schlenning W.D., Bode W.;  
RT "Catalytic domain structure of vampire bat plasminogen activator: a  
RT molecular paradigm for proteolysis without activation cleavage.";  
RL Biochemistry 36:13483-13493(1997).  
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS  
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-val bond in  
CC plasminogen to form plasmin.  
CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN  
CC THE PRESENCE OF FIBRIN I.  
CC -1- SUBUNIT: MONOMER.  
CC -1- DOMAIN: THE FIBRONECTIN TYPE-1 DOMAIN MEDIATES BINDING TO FIBRIN,  
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED  
CC STIMULATION OF ACTIVITY.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
-----  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: M63987; AAA31591.1; -;  
DR EMBL: M63986; AAA31592.1; -;  
DR PDB: 1A5T; 23-MAR-99.  
DR MEROPS: S01.232; -;  
DR GlycoSuiteDB: P98119; -;  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000083; Fibrinectn.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR001254; Ser\_protease-Try.  
DR Pfam: PF000008; EGF; 1.  
DR Pfam: PF000039; fn1; 1.  
DR Pfam: PF000051; kringle; 1.  
DR Pfam: PF000089; trypsin; 1.  
DR PRINTS: PRO00722; CHYMOTRYPSIN.  
DR PRINTS: PRO00018; KRINGLE.  
DR ProDom: PD000395; Kringle; 1.  
DR SMART: SM00181; EGF; 1.  
DR SMART: SM00058; FN1; 1.  
DR SMART: SM00130; KR; 1.  
DR SMART: SM00020; Tryp\_Spc; 1.

DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.  
 DR PROSITE: PS00070; KRINGLE\_2; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KM Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KRINGLE; EGF-like domain; Signal; Multigene family; 3D-structure.  
 FT SIGNAL 1 36  
 FT CHAIN 37 477  
 FT DOMAIN 40 82  
 FT DOMAIN 83 121  
 FT DOMAIN 128 209  
 FT ACT\_SITE 225 477  
 FT ACT\_SITE 272 272  
 FT ACT\_SITE 321 321  
 FT ACT\_SITE 428 428  
 FT DISULFID 42 72  
 FT DISULFID 70 79  
 FT DISULFID 87 98  
 FT DISULFID 92 109  
 FT DISULFID 111 120  
 FT DISULFID 128 209  
 FT DISULFID 149 191  
 FT DISULFID 180 204  
 FT DISULFID 214 345  
 FT DISULFID 257 272  
 FT DISULFID 265 334  
 FT DISULFID 359 434  
 FT DISULFID 391 407  
 FT DISULFID 424 452  
 FT CARBOHYD 153 153  
 FT CARBOHYD 398 398  
 SQ SEQUENCE 477 AA; 53616 MW; 53616 MW; AAO6FD1739C10E5E CRC64;  
 Query Match 41.7%; Score 208; DB 1; Length 477;  
 Best Local Similarity 45.2%; Pred. No. 1.2e-16;  
 Matches 38; Conservative 11; Mismatches 35; Indels 0; Gaps 0;  
 QY 2 TCEGNGHFYRKASTDTMGRCPLPWNASATVLOQTYHAHRSNALQGLGKHNYCRPNDR 61  
 DB 127 TCEGNGGVYTRGTWSTAESRVEECINNNSSLLTRTYNGMPAFNGLGNHNYCRPNCA 186  
 QY 62 RRPWCYVOYGLKRLVOECVHICA 85  
 DB 187 PRWCYVIRAKRFTSESCVPCS 210  
 RESULT 15  
 UROK\_CHICK STANDARD; PRT; 434 AA.  
 ID UROK\_CHICK  
 AC P15120;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)  
 DE (U-plasminogen activator).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90110185; PubMed=2295632;  
 RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;  
 RT "The chicken urokinase-type plasminogen activator gene."  
 RT J. Biol. Chem. 265:1339-1344(1990).  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in

CC plasminogen to form plasmin.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J05187; AAA49131.1; -  
 CC EMBL: J05188; AAA49130.1; -  
 CC PIR: A35005; A35005.  
 CC HSSP: P00763; IDPO.  
 CC MEROPS: S01.231; -  
 CC InterPro: IPR001314; Chymotrypsin.  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR000001; Kringle.  
 CC InterPro: IPR001254; Ser protease\_Try.  
 CC Pfam: PF00051; kringle\_1.  
 CC Pfam: PF00089; kringle\_1.  
 CC PRINTS: PR00722; TRYPSIN.  
 CC ProDom: PD000395; kringle\_1.  
 CC SMART: SM00181; EGF\_1.  
 CC SMART: SM00130; TRYP\_SPE; 1.  
 DR SKART: SM00020; TRYP\_SPE; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.  
 DR PROSITE: PS00070; KRINGLE\_2; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW KRINGLE; EGF-like domain; Signal; Multigene family; 3D-structure.  
 FT SIGNAL 1 36  
 FT CHAIN 37 477  
 FT DOMAIN 40 82  
 FT DOMAIN 83 121  
 FT DOMAIN 128 209  
 FT ACT\_SITE 225 477  
 FT ACT\_SITE 272 272  
 FT ACT\_SITE 321 321  
 FT ACT\_SITE 428 428  
 FT DISULFID 42 72  
 FT DISULFID 70 79  
 FT DISULFID 87 98  
 FT DISULFID 92 109  
 FT DISULFID 111 120  
 FT DISULFID 128 209  
 FT DISULFID 149 191  
 FT DISULFID 180 204  
 FT DISULFID 214 345  
 FT DISULFID 257 272  
 FT DISULFID 265 334  
 FT DISULFID 359 434  
 FT DISULFID 391 407  
 FT DISULFID 424 452  
 FT CARBOHYD 153 153  
 FT CARBOHYD 398 398  
 SQ SEQUENCE 477 AA; 53616 MW; 53616 MW; AAO6FD1739C10E5E CRC64;  
 Query Match 40.8%; Score 203.5; DB 1; Length 434;  
 Best Local Similarity 55.9%; Pred. No. 3.5e-16;  
 Matches 38; Conservative 6; Mismatches 19; Indels 5; Gaps 2;  
 QY 3 CYEGNGHFYRKASTDTMGRCPLPWNASATVLOQTYHAHRSNALQGLGKHNYCRPNDR 61  
 DB 79 CYSGNGEDYRKMAEDP-----GLTIWDHPVYIRMGDYHADIKNALQGLGKHNYCRPNDR 134  
 QY 62 RRPWCYVOY 69  
 DB 135 SRPWCYTK 142

Tue Nov 12 18:06:47 2002

pct-us02-27855-1.rsp

Page 15

Search completed: November 11, 2002, 13:00:25  
Job time : 22 secs

---

•  
•  
•  
•



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 12:24:41 ; Search time 75 Seconds  
(without alignments)  
236.267 Million cell updates/sec

Title: PCT-US02-27855-1

Perfect score: 499

Sequence: 1 KTCYEGNGHFYRGAASDTM.....YVQGLKPLVGECAVHDCAD 86

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPREMBL\_21:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.rodent:\*
- 12: sp.virus:\*
- 13: sp.vertibrate:\*
- 14: sp.unclassified:\*
- 15: sp.virus:\*
- 16: sp.bacteriophage:\*
- 17: sp.archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	97.0	154	4 Q96SE8	Q96SE8 homo sapien
2	383	76.8	157	6 Q9TVA8	Q9TVA8 bos taurus
3	306	61.3	214	6 Q9XT70	Q9XT70 oryctolagus
4	217	43.5	559	11 Q91VP2	Q91VP2 mus musculus
5	216	43.3	395	4 Q9BZM1	Q9BZM1 homo sapien
6	216	43.3	516	4 Q9BQ99	Q9BQ99 homo sapien
7	211	42.3	562	6 Q8S023	Q8S023 sus scrofa
8	191	38.3	202	13 Q90675	Q90675 gallus gall
9	188.5	37.8	560	4 Q14520	Q14520 homo sapien
10	188	37.7	653	11 Q8VCS4	Q8VCS4 mus musculus
11	186.5	37.4	597	11 Q35727	Q35727 mus musculus
12	177.5	35.6	616	6 Q97507	Q97507 sus scrofa
13	157	31.5	452	13 Q90Y90	Q90Y90 xenopus lae
14	153.5	30.8	812	11 Q9R0W3	Q9R0W3 rattus norv
15	153	30.7	454	6 Q46506	Q46506 papio hamad
16	153	30.7	806	6 Q18783	Q18783 macropus eu

17	152	30.5	113	4 Q9UR6	Q9UR6 homo sapien
18	151.5	30.4	359	6 Q8MMR1	Q8MMR1 canis faml
19	151	30.3	113	4 Q9UR7	Q9UR7 homo sapien
20	151	30.3	113	4 Q9UR5	Q9UR5 homo sapien
21	149	29.9	810	4 Q15146	Q15146 homo sapien
22	148.5	29.8	709	13 Q902N6	Q902N6 brachydanto
23	147.5	29.6	399	4 Q96GL8	Q96GL8 homo sapien
24	147.5	29.6	420	4 Q9BTP9	Q9BTP9 homo sapien
25	147.5	29.6	716	13 Q91691	Q91691 xenopus lae
26	147	29.5	393	4 Q9BRB6	Q9BRB6 homo sapien
27	146	29.3	105	4 Q9UR8	Q9UR8 homo sapien
28	145	29.1	648	4 Q9H1V4	Q9H1V4 homo sapien
29	144	28.9	567	4 Q13208	Q13208 homo sapien
30	142	28.5	902	5 Q17576	Q17576 caenorhabdl
31	142	28.5	928	5 Q9BLV1	Q9BLV1 caenorhabdl
32	139.5	28.0	717	13 P70006	P70006 xenopus lae
33	138	27.7	1145	5 Q9BKJ8	Q9BKJ8 aplysia cal
34	137.5	27.6	704	13 Q90865	Q90865 gallus gall
35	137	27.5	812	11 Q91WJ5	Q91WJ5 mus musculu
36	135	27.1	132	4 Q16609	Q16609 mus musculu
37	135	27.1	145	6 Q28911	Q28911 macaca fasc
38	134	26.9	685	5 Q24488	Q24488 drosophila
39	133.5	26.8	313	13 Q9PU78	Q9PU78 crocodylus
40	132	26.5	263	4 Q00318	Q00318 homo sapien
41	132	26.5	263	4 Q96FE7	Q96FE7 homo sapien
42	131.5	26.4	215	13 Q42341	Q42341 gallus gall
43	131.5	26.4	716	11 Q91XG8	Q91XG8 mus musculu
44	131.5	26.4	726	13 Q90978	Q90978 gallus gall
45	131.5	26.4	2869	6 Q28398	Q28398 erinaceus e

## ALIGNMENTS

RESULT 1

ID Q96SE8 PRELIMINARY; PRT; 154 AA.

AC Q96SE8;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Urokinase-type plasminogen activator amino-terminal fragment.

DE ATF.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

XX [1]

RP SEQUENCE FROM N.A.

RA Fu J., Bai X., Ruan C.;

RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator.";

RT Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.

RL [2]

RP SEQUENCE FROM N.A.

RA Bai X., Fu J., Wang W., Xi X., Ruan C.;

RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis.";

RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.

EMBL: AV029537; AAK38734.1; -

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR000001; Kringle.

DR Pfam: PF00051; Kringle; 1.

DR ProDom: PDOM00395; Kringle; 1.

DR SMART: SM00181; EGF; 1.

DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.

DR PROSITE: PS00021; UNKNOWN\_1.

DR PROSITE: PS00070; KRINGLE\_2; 1.

KW kinase.

SO SEQUENCE 154 AA; 17305 MW; A3CCF2PCFF505572 CRC64;

Query Match 97.0%; Score 484; DB 4; Length 154;  
 Best Local Similarity 97.7%; Pred. No. 3.2e-52;  
 Matches 84; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASDTMGRCPLPWSATYVLOQTYHAHRSNALQLGKHNKCRNPN 60  
 |||||  
 DB 68 KTCYEGNGHFYRGKASDTMGRCPLPWSATYVLOQTYHAHRSNALQLGKHNKCRNPN 127  
 |||||  
 OY 61 RRRPCYVQVGLKPLVQECVHDCAD 86  
 |||||  
 DB 128 RRRPCYVQVGLKPLVQECVHDCAD 153  
 |||||

## RESULT 2

O9TVAB PRELIMINARY; PRT; 157 AA.  
 AC O9TVAB;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Urokinase plasminogen activator (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 NCBI\_Taxid=9913;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SKELETAL MUSCLE;  
 RA MEDLINE-21071388; PubMed-11204721;  
 RA Balcerzak D. Quereguesser L., Dixon W.T., Baracos V.E.,  
 RT "Coordinate expression of matrix-degrading proteinases and their  
 RT activators and inhibitors in bovine skeletal muscle."  
 RL J Anim Sci 79:94-107(2001).  
 DR EMBL: AF144761; AAD30301.1; -.  
 DR HSSP: P00749; 1.0RK.  
 DR MEROPS: S01.231; -.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000001; Kringle.  
 DR Pfam: PF00051; Kringle.1.  
 DR PRINTS: PR00018; KRINGLE.1.  
 DR ProDom: PD000395; Kringle; 1.  
 DR SMART: SM00130; KR; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.  
 DR PROSITE: PS00070; KRINGLE\_2; 1.  
 DR Kinase.  
 DR KINASE.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 157 AA; 17858 MW; A768DC72C1BFB7 CRC64;

Query Match 76.8%; Score 383; DB 6; Length 157;  
 Best Local Similarity 75.3%; Pred. No. 1.1e-35;  
 Matches 64; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASDTMGRCPLPWSATYVLOQTYHAHRSNALQLGKHNKCRNPN 60  
 |||||  
 DB 36 KTCYEGNGHFYRGKASDTMGRCPLPWSATYVLOQTYHAHRSNALQLGKHNKCRNPN 95  
 |||||  
 OY 61 RRRPCYVQVGLKPLVQECVHDCAD 85  
 |||||  
 DB 96 RRRPCYVQVGLKPLVQECVHDCAD 120  
 |||||

## RESULT 3

O9XT70 PRELIMINARY; PRT; 214 AA.  
 AC O9XT70;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Urokinase-type plasminogen activator (Fragment).  
 OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NCBI\_Taxid=9986;  
 [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE-LUNG;  
 RA Yin J., Igell S.,  
 RT "Partial mRNA of rabbit uPA."  
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL: AF097647; AAD39351.1; -.  
 DR HSSP: P00749; 1.0RN.  
 DR MEROPS: S01.231; -.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000001; Kringle.  
 DR Pfam: PF00051; Kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR ProDom: PD000395; Kringle; 1.  
 DR SMART: SM00130; KR; 1.  
 DR SMART: SM00020; TRYPSIN; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.  
 DR PROSITE: PS00070; KRINGLE\_2; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR Hydroxylase; Kinase; Serine protease.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 214 AA; 24314 MW; 69975CA1C32B07E CRC64;

Query Match 61.3%; Score 306; DB 6; Length 214;  
 Best Local Similarity 81.2%; Pred. No. 5.6e-30;  
 Matches 52; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 20 MGRPCLPWSATYVLOQTYHAHRSNALQLGKHNKCRNPNRRPCYVQVGLKPLVQEC 79  
 |||||  
 DB 1 MDRPCLPWSATYVLOQTYHAHRSNALQLGKHNKCRNPNRRPCYVQVGLKPLVQEC 60  
 |||||  
 OY 80 MVHD 83  
 |||||  
 DB 61 KVHD 64

## RESULT 4

O91VP2 PRELIMINARY; PRT; 559 AA.  
 AC O91VP2;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Similar to plasminogen activator, tissue.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAUSBERG R.;  
 RA Strausberg R.;  
 RL EMBL: BC011256; AAH11256.1; -.  
 DR EMBL: BC011256; AAH11256.1; -.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000083; Fibrinogen.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR001254; Ser-protease-TRY.  
 DR Pfam: PF00089; EGF; 1.  
 DR Pfam: PF00039; EGF; 1.  
 DR Pfam: PF00051; Kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR ProDom: PD000395; Kringle; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.

DR PROSITE: PS01186; EGF\_2; UNKNOWN\_1.  
DR PROSITE: PS01253; FIBRONECTIN\_1; UNKNOWN\_1.  
DR PROSITE: PS00021; KRINGLE\_1; UNKNOWN\_2.  
DR PROSITE: PS00070; KRINGLE\_2; 2.  
DR PROSITE: PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPSIN\_SER; UNKNOWN\_1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;

Query Match 43.5%; Score 217; DB 11; Length 559;  
Best Local Similarity 47.0%; Pred. No. 1.8e-18;  
Matches 39; Conservative 11; Mismatches 33; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWNASATVLTQTYHAHRSNALQLGKHNKCRPNDR 62  
DB 213 CYVGKCVYTRNGHSLTSGASCLPWNSTYLMGKSTYAWRTNSQALGLGRHNCRNPDGA 272  
QY 63 RPKCVYVGLKPLVQECMVHDC 85  
DB 273 RPKCHVAKNRRLTWECVDPSCS 295

RESULT 5  
Q9B2M1 PRELIMINARY; PRT; 395 AA.

AC 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Neontal thrombolytic agent alpha-form (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dou D.;  
RT "A brain-type plasminogen activator";  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
TRYPSIN FAMILY.  
DR EMBL: AF260825; AAK11956.1; -.  
DR HSP: P00750; 1PK2.  
DR MEROPS: S01.232; -.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000083; Fibrinctn.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR001254; Ser-protease\_Try.  
DR Pfam: PF00039; fnl; 1.  
DR Pfam: PF00051; kringle; 1.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR00018; KRINGLE.  
DR ProDom: PD000395; Kringle; 1.  
DR SMART: SM00058; FNL; 1.  
DR SMART: SM00130; KR; 1.  
DR SMART: SM00020; TRYP\_SPE; 1.  
DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE: PS00021; KRINGLE\_1; UNKNOWN\_1.  
DR PROSITE: PS50070; KRINGLE\_2; 1.  
DR PROSITE: PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
FT NON\_TER 395  
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 43.3%; Score 216; DB 4; Length 395;  
Best Local Similarity 47.0%; Pred. No. 1.6e-18;  
Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;  
QY 3 CYEGNGHFYRGKASTDTMGRCPLPWNASATVLTQTYHAHRSNALQLGKHNKCRPNDR 62

DB 48 CYFGNGSYRRTHTSLTEGASCLPWNSTYLMGKSTYAWRTNSQALGLGRHNCRNPDGA 107  
QY 63 RPKCVYVGLKPLVQECMVHDC 85  
DB 108 RPKCHVAKNRRLTWECVDPSCS 130

RESULT 6  
Q9B099 PRELIMINARY; PRT; 516 AA.

AC 09B099;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Similar to plasminogen activator, tissue.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
TRYPSIN FAMILY.  
DR EMBL: BC002795; AA02795.1; -.  
DR HSP: P00750; 1ASH.  
DR MEROPS: S01.232; -.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR001254; Ser-protease\_Try.  
DR Pfam: PF00008; EGF; 1.  
DR Pfam: PF00051; kringle; 2.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR00018; KRINGLE.  
DR ProDom: PD000395; Kringle; 2.  
DR SMART: SM00181; EGF-like; 1.  
DR SMART: SM00001; EGF-like; 1.  
DR SMART: SM00130; KR; 2.  
DR SMART: SM00020; TRYP\_SPE; 1.  
DR PROSITE: PS00022; EGF\_2; 1.  
DR PROSITE: PS01186; EGF\_2; 1.  
DR PROSITE: PS00021; KRINGLE\_1; UNKNOWN\_2.  
DR PROSITE: PS50070; KRINGLE\_2; 2.  
DR PROSITE: PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW EGF-like domain; glycoprotein; Hydrolase; Serine protease.  
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FC96800 CRC64;

Query Match 43.3%; Score 216; DB 4; Length 516;  
Best Local Similarity 47.0%; Pred. No. 2.2e-18;  
Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWNASATVLTQTYHAHRSNALQLGKHNKCRPNDR 62  
DB 169 CYFGNGSYRRTHTSLTEGASCLPWNSTYLMGKSTYAWRTNSQALGLGRHNCRNPDGA 228  
QY 63 RPKCVYVGLKPLVQECMVHDC 85  
DB 229 RPKCHVAKNRRLTWECVDPSCS 251

RESULT 7  
Q8SQ23 PRELIMINARY; PRT; 562 AA.  
AC 08SQ23;  
DT 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE T-plasminogen activator.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae, Sus.  
 RN NCBI\_TaxID=9823;  
 RP  
 RC  
 RA  
 RT  
 RL  
 DR  
 SQ

Query Match 42.3%; Score 211; DB 6; Length 562;  
 Best Local Similarity 48.8%; Pred. No. 1e-17;  
 Matches 40; Conservative 7; Mismatches 35; Indels 0; Gaps 0;

QY 3 CYEGNGHYFGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNDNR 62  
 DB 215 CYTGKGLDYRGTRTSLSGAFCLPWNSTLVMGKITAMNSNAQTGLGKHNCRNDNR 274  
 QY 63 RPMCYYOVGLKPLVQECMVHDC 84  
 DB 275 QPWCYVGLKDKHKLWECVDCDLPQC 296

RESULT 8  
 ID 090675 PRELIMINARY; PRT; 202 AA.  
 AC 090675;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Tissue-type plasminogen activator (Fragment).  
 GN TPA.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 RN NCBI\_TaxID=9031;  
 RP  
 RC  
 RA  
 RT  
 RL  
 DR  
 SQ

Query Match 38.3%; Score 191; DB 13; Length 202;  
 Best Local Similarity 43.9%; Pred. No. 9.7e-16;  
 Matches 36; Conservative 7; Mismatches 39; Indels 0; Gaps 0;

QY 3 CYEGNGHYFGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNDNR 62  
 DB 41 CYTGNLAIRGYSRTRKSGFSLPWNVPVLSKITYALREORRALGLGKHNCRNDNR 100  
 QY 63 RPMCYYOVGLKPLVQECMVHDC 84  
 DB 101 QPWCYVGLKDKHKLWECVDCDLPQC 122

RESULT 9  
 ID 014520 PRELIMINARY; PRT; 560 AA.  
 AC 014520; 000663;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE HGF activator like protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN NCBI\_TaxID=9606;  
 RP  
 RC  
 RA  
 RT  
 RL  
 DR  
 SQ

Query Match 37.8%; Score 188.5; DB 4; Length 560;  
 Best Local Similarity 40.5%; Pred. No. 6.2e-15;  
 Matches 34; Conservative 16; Mismatches 33; Indels 1; Gaps 1;

QY 3 CYEGNGHYFGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNDNR 62  
 DB 194 CYVGDDYSYRGKNNRTVNOHACLYWNSHLLQENTVMFWMEDAEHTGIGHNCRNDNR 253  
 QY 63 RPMCYYOVGLKPLVQECMVHDC 85

Db 254 KPWCFIKVTNDKWKECVDSACS 277

## RESULT 10

08VCS4 PRELIMINARY; PRT; 653 AA.

AC 08VCS4;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Hypothetical 70.6 kDa protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC019376; AAH19376.1; -  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001438; EGF-II.  
 DR InterPro: IPR000083; Fibrinctn1.  
 DR InterPro: IPR000562; FN\_Type\_II.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00008; EGF\_2.  
 DR Pfam: PF00039; fn1; 1.  
 DR Pfam: PF00040; fn2; 1.  
 DR Pfam: PF00051; kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PRO00722; CHYMOTRYPSIN.  
 DR PRINTS: PRO0010; EGFBLD.  
 DR PRINTS: PRO0013; FNYPEIT.  
 DR PRINTS: PRO0018; KRINGLE.  
 DR ProDom: PD000395; kringle; 1.  
 DR ProDom: PD000995; FN\_Type\_II; 1.  
 DR SMART: SM00181; EGF\_2.  
 DR SMART: SM00058; FN1; 1.  
 DR SMART: SM00059; FN2; 1.  
 DR SMART: SM00130; KR; 1.  
 DR SMART: SM0020; Tryp\_Spc; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_2.  
 DR PROSITE: PS01186; EGF\_2; UNKNOWN\_1.  
 DR PROSITE: PS01253; FIBRONECTIN\_1; UNKNOWN\_1.  
 DR PROSITE: PS00023; FIBRONECTIN\_2; UNKNOWN\_1.  
 DR PROSITE: PS00021; KRINGLE\_1; UNKNOWN\_1.  
 DR PROSITE: PS00070; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; UNKNOWN\_1.  
 KW Hypothetical protein. 653 AA; 70553 MW; FE18D90174EDFDD CRC64;  
 SQ

Query Match 37.7%; Score 188; DB 11; Length 653;  
 Best Local Similarity 53.0%; Pred. No. 8, 5e-15;  
 Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 3 CYEGNGHFRGKASTDTMGRCPLPMSATVLOQTYHAHNSNALQLGLGKHNCRNDNR 62  
 DB 283 CFLNGTEYRGVASTAASGLSCLAWNSDLLYQELHVDVAVAALLGLGPHAYCRNDKDE 342

QY 63 RPKCYV 68  
 DB 343 RPKCYV 348

RESULT 11  
 035727

ID 035727 PRELIMINARY; PRT; 597 AA.

AC 035727;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Factor XII.  
 GN F12.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Schloesser M., Schwager S., Engel W.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.  
 DR EMBL: X99571; CA67891.1; -  
 DR HSP: P00760; 1A07.  
 DR MEROPS: S01.211; -  
 DR MGD: MGI:1891012; F12.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000083; Fibrinctn1.  
 DR InterPro: IPR000562; FN\_Type\_II.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00008; EGF\_2.  
 DR Pfam: PF00039; fn1; 1.  
 DR Pfam: PF00040; fn2; 1.  
 DR Pfam: PF00051; kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PRO00722; CHYMOTRYPSIN.  
 DR PRINTS: PRO0018; KRINGLE.  
 DR ProDom: PD000395; kringle; 1.  
 DR ProDom: PD000995; FN\_Type\_II; 1.  
 DR SMART: SM00181; EGF\_2.  
 DR SMART: SM00058; FN1; 1.  
 DR SMART: SM00059; FN2; 1.  
 DR SMART: SM00130; KR; 1.  
 DR SMART: SM0020; Tryp\_Spc; 1.  
 DR PROSITE: PS00022; EGF\_1; 2.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE: PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.  
 DR PROSITE: PS50070; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW EGF-like domain; Glycoprotein; Hydrolyase; Serine protease.  
 SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 37.4%; Score 186.5; DB 11; Length 597;  
 Best Local Similarity 44.0%; Pred. No. 1, 2e-14;  
 Matches 37; Conservative 8; Mismatches 34; Indels 5; Gaps 2;

QY 2 TCYEGNGHFRGKASTDTMGRCPLPMSATVLOQTYHAHNSNALQLGLGKHNCRNDNR 60  
 DB 216 TCYEGRLSTYRGAGTQSGAPCORW---TVATRYNMTKQALSWGLGHAFCRNDPN 271

QY 61 RRPFCYVGVGLKPLVQECMVHDC 84  
 DB 272 DTRPFCYVSGDRLSMDYCGLEDC 295

RESULT 12  
 097507  
 ID 097507 PRELIMINARY; PRT; 616 AA.  
 AC 097507;  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 OS Euk. scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 RN NCB1\_TaxID=9823;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Takahashi T., Kihara T.;  
 RT "Porcine liver factor XII."  
 RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL: A022426; BAA37148.1; -.  
 DR HSSP: P00763; IDPO.  
 DR MEROPS: S01.211; -.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR000083; Fibrinectn1.  
 DR InterPro: IPR000562; FN\_Type\_II.  
 DR InterPro: IPR000001; Kringle.  
 DR Pfam: PF00008; EGF\_2.  
 DR Pfam: PF00039; fn1; 1.  
 DR Pfam: PF00040; fn2; 1.  
 DR Pfam: PF00051; kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00013; ENTYPERII.  
 DR PRINTS: PR00018; KRINGLE.  
 DR ProDom: PD000995; kringle; 1.  
 DR ProDom: PD000995; FN\_Type\_II; 1.  
 DR SMART: SM00181; EGF\_2.  
 DR SMART: SM00058; FN1; 1.  
 DR SMART: SM00059; FN2; 1.  
 DR SMART: SM00130; KR; 1.  
 DR SMART: SM00020; TRYP-spec; 1.  
 DR PROSITE: PS00022; EGF\_1; 2.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE: PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.  
 DR PROSITE: PS00070; KRINGLE\_2; 1.  
 DR PROSITE: PS00240; TRYPsin\_DOM; 1.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 DR EGF-like domain; glycoprotein; Hydrolase; Serine protease.  
 SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 35.6%; Score 177.5; DB 6; Length 616;  
 Best Local Similarity 44.6%; Pred. No. 1.6e-13;  
 Matches 37; Conservative 6; Mismatches 35; Indels 5; Gaps 2;

QY 3 CYENGHRYRGKASTDWM--GRPCLPWNSATVYQOYHAHRSNALQLGKHNCRPNDR 61  
 DB 217 CYSRGLSTYRKMAQTITSGAPCPWAS---EATYNNMTAEQALNMGIDGHACRPNDND 272  
 QY 62 RRPWCYQVGLKPLVQECWVHDC 84  
 DB 273 TRPWCYWRGDQLSMQYCRILARC 295

RESULT 13  
 ID Q90Y90 PRELIMINARY; PRT; 452 AA.  
 AC Q90Y90;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE KREMEN.  
 DR KREMEN.  
 DR InterPro: IPR000001; Chymotrypsin.

OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 RN NCB1\_TaxID=8355;  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-21167372; PubMed-1126760;  
 RA Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,  
 RA Nakamura T., Nakamura T.;  
 RT "Molecular cloning and characterization of kremen, a novel kringle-  
 RT containing transmembrane protein."  
 RL Blochm. Biophys. Acta 1518:63-72(2001).  
 DR EMBL: A070851; BAB64284.1; -.  
 DR InterPro: IPR000859; CUB\_domain.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR002889; WSC.  
 DR Pfam: PF00431; CUB; 1.  
 DR Pfam: PF01822; WSC; 1.  
 DR Pfam: PF00051; kringle; 1.  
 DR ProDom: PD000395; kringle; 1.  
 DR PROSITE: PS01180; CUB; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; UNKNOWN\_1.  
 DR PROSITE: PS50070; KRINGLE\_2; 1.  
 SQ SEQUENCE 452 AA; 50188 MW; ED24BCD1AF4564E2 CRC64;

Query Match 31.5%; Score 157; DB 13; Length 452;  
 Best Local Similarity 44.1%; Pred. No. 3.9e-11;  
 Matches 30; Conservative 9; Mismatches 25; Indels 4; Gaps 2;

QY 3 CYENGHRYRGKASTDWM--GRPCLPWNSATVYQOYHAHRSNALQLGKHNCRPNDR 60  
 DB 30 CYTVNGADYRGQTQNTSLDGKRCPLFWNE--TRQHPYNTLKRYNGSGIGENHYCNRPDC 87  
 QY 61 RRRPWCYV 68  
 DB 88 DVSPWCYI 95

RESULT 14  
 ID Q9R0W3 PRELIMINARY; PRT; 812 AA.  
 AC Q9R0W3;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Plasminogen protein precursor (EC 3.4.21.7).  
 GN PLASMINOGEN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN NCB1\_TaxID=10116;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Bangert K., Johnsen A.H., Thorsen S.;  
 RT "Rat plasminogen: cDNA and gene structure."  
 RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE-91250378; PubMed-1645711;  
 RA Kanalas J.T., Makler S.P.;  
 RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a  
 RT receptor site for plasminogen."  
 RL J. Biol. Chem. 266:10823-10829(1991).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL: A0242649; CAB6014.1; -.  
 DR HSSP: P00747; IPWK.  
 DR MEROPS: S01.233; -.  
 DR InterPro: IPR001314; Chymotrypsin.

DR InterPro: IPR003014; PAN.  
DR InterPro: IPR003609; Pan\_app.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR InterPro: IPR001400; Somatotropin.  
DR Pfam: PF00051; kringle; 5.  
DR Pfam: PF00024; PAN; 1.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR00018; KRINGLE.  
DR ProDom: PD000395; kringle; 5.  
DR SMART: SM00130; KR; 4.  
DR SMART: SM00473; PAN\_Ap; 1.  
DR SMART: SM00020; Tryp\_Spc; 1.  
DR PROSITE: PS00021; KRINGLE\_1; 5.  
DR PROSITE: PS50070; KRINGLE\_2; 5.  
DR PROSITE: PS50240; SOMATOTROPIN\_2; UNKNOWN\_1.  
DR PROSITE: PS50240; TRYPsin\_DOM; 1.  
DR PROSITE: PS00134; TRYPsin\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
KW Hydrolyase; Serine protease; Signal.  
FT SIGNAL 1 19  
FT CHAIN 1 20  
SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

Query Match 30.8%; Score 153.5; DB 11; Length 812;  
Best Local Similarity 38.9%; Pred. No. 2.1e-10;  
Matches 35; Conservative 13; Mismatches 27; Indels 15; Gaps 6;

OY 3 CYEGNGHFYRGKASTDTMGRCPLPWNASATVLOQTYHAHR---SNALQGLGKHNYCRNPD 59  
DB 376 CYGNGKSTYRGSTSTNTGKQCSW---VSMTPSHSKTPANPDAGL-EMNYCRNPD 429  
OY 60 N-RRRWCYVQGLKPLY--QECMVHDCAD 86  
DB 430 NDGRGWCFT---TPSVRWEYCNLRCSE 456

RESULT 15  
O46506 PRELIMINARY; PRT; 454 AA.  
AC O46506.  
DT 01-JUN-1998 (TRENBLREL. 06, Created)  
DT 01-JUN-1998 (TRENBLREL. 06, Last sequence update)  
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
DE Apolipoprotein a (Fragment).  
GN BABAPOA.  
OS Papio hamadryas (Hamadryas baboon).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Papio.  
OX NCBI\_TaxID=9557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cox L.A., Jett C., Hixson J.E.;  
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice  
RT Site Mutation is Associated with Deletion of a Single Exon in a Null  
RT Allele."  
RU Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR EMBL: AF029691; AAB97886.1; -.  
DR HSSP: P00747; 2PK4.  
DR MEROPS: S01.226; -.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00051; kringle; 2.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR00018; KRINGLE.  
DR ProDom: PD000395; Kringle; 2.  
DR SMART: SM00130; KR; 2.  
DR SMART: SM00020; Tryp\_Spc; 1.

DR PROSITE: PS00021; KRINGLE\_1; 2.  
DR PROSITE: PS50070; KRINGLE\_2; 2.  
DR PROSITE: PS50240; TRYPsin\_DOM; 1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
KW Hydrolyase; Lipoprotein; Serine protease.  
FT NON\_TER 1 1  
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 30.7%; Score 153; DB 6; Length 454;  
Best Local Similarity 38.6%; Pred. No. 1.2e-10;  
Matches 34; Conservative 11; Mismatches 31; Indels 12; Gaps 4;

OY 3 CYEGNGHFYRGKASTDTMGRCPLPWNASATVLO--QTYHAHRSNALQGLGKHNYCRNPD 60  
DB 102 CYHDDGQSYRGSFSTYVGTGRCQSSMTPHQHKRTEPNPNGLTW-----NYCRNPD 156  
OY 61 RRRWCYVQGLKPLY--QECMVHDCAD 86  
DB 157 DTGFWCFT---MDPSVRWEYCNLRCSD 181

Search completed: November 11, 2002, 13:01:56  
Job time : 78 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: November 11, 2002, 11:35:58 ; Search time 2622 Seconds  
(without alignments)  
2863.663 Million cell updates/sec

Title: PCT-US02-27855-2

Perfect score: 258  
Sequence: 1 aaacctgctatgaggggaa.....tggctgcatgactgcagat 258

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml: \*  
1: gb\_da: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pac: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vtl: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	256.4	99.4	423	6	A76865	A76865 Sequence 3
2	256.4	99.4	1236	6	E02578	E02578 DNA encodin
3	256.4	99.4	1236	6	E02579	E02579 Human prot
4	256.4	99.4	1236	6	E02708	E02708 DNA sequenc
5	256.4	99.4	1236	6	E02709	E02709 DNA sequenc
6	256.4	99.4	1236	6	E02710	E02710 DNA sequenc
7	256.4	99.4	1236	6	E02832	E02832 DNA sequenc
8	256.4	99.4	1236	6	E02833	E02833 DNA encodin
9	256.4	99.4	1236	6	E02577	E02577 DNA encodin
10	256.4	99.4	1236	6	E02647	E02647 DNA sequenc
11	256.4	99.4	1236	6	E02649	E02649 DNA sequenc
12	256.4	99.4	1236	6	E02711	E02711 DNA sequenc
13	256.4	99.4	1236	6	E06064	E06064 DNA encodin
14	256.4	99.4	1359	6	AX451990	AX451990 Sequence
15	256.4	99.4	1394	6	E00421	E00421 cDNA coding
16	256.4	99.4	1473	6	E02114	E02114 cDNA sequen
17	256.4	99.4	1474	6	E00924	E00924 cDNA encodi
18	256.4	99.4	1474	6	E01559	E01559 cDNA sequen
19	256.4	99.4	1474	6	E01580	E01580 cDNA encodi
20	256.4	99.4	1474	6	E02095	E02095 DNA sequenc
21	256.4	99.4	1475	6	AX365729	AX365729 Sequence
22	256.4	99.4	1475	6	E01560	E01560 cDNA sequen
23	256.4	99.4	1475	9	H000001	H000001 Human prepr
24	256.4	99.4	1475	9	H000001	H000001 Human prepr
25	256.4	99.4	1475	9	H000001	H000001 Human prepr
26	256.4	99.4	1475	9	H000001	H000001 Human prepr
27	256.4	99.4	1475	9	H000001	H000001 Human prepr
28	256.4	99.4	1475	9	H000001	H000001 Human prepr
29	256.4	99.4	1475	9	H000001	H000001 Human prepr
30	256.4	99.4	1475	9	H000001	H000001 Human prepr
31	256.4	99.4	1475	9	H000001	H000001 Human prepr
32	256.4	99.4	1475	9	H000001	H000001 Human prepr
33	256.4	99.4	1475	9	H000001	H000001 Human prepr
34	256.4	99.4	1475	9	H000001	H000001 Human prepr
35	256.4	99.4	1475	9	H000001	H000001 Human prepr
36	256.4	99.4	1475	9	H000001	H000001 Human prepr
37	256.4	99.4	1475	9	H000001	H000001 Human prepr
38	256.4	99.4	1475	9	H000001	H000001 Human prepr
39	256.4	99.4	1475	9	H000001	H000001 Human prepr
40	256.4	99.4	1475	9	H000001	H000001 Human prepr
41	256.4	99.4	1475	9	H000001	H000001 Human prepr
42	256.4	99.4	1475	9	H000001	H000001 Human prepr
43	256.4	99.4	1475	9	H000001	H000001 Human prepr
44	256.4	99.4	1475	9	H000001	H000001 Human prepr
45	256.4	99.4	1475	9	H000001	H000001 Human prepr

#### ALIGNMENTS

RESULT 1  
A76865  
LOCUS A76865 423 bp DNA linear PAT 19-OCT-1999  
DEFINITION Sequence 3 from Patent WO9315199.  
ACCESSION A76865  
VERSION A76865.1 GI:6088666  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
AUTHORS Flier, R. and Fournier, A.  
TITLE NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, PREPARATION THEREOF AND  
JOURNAL PHARMACEUTICAL COMPOSITION CONTAINING SAID POLYPEPTIDES  
Patent: WO 9315199-A 3 05-AUG-1993;

Pred. No. is the number of results predicted by chance to have a

FEATURES	PHONE	POULENC	RORER SA (FR)	Location/Qualifiers
SOURCE	1..423			/organism="unidentified"
	1..423			/db_xref="taxon:32644"
CDS	<3..419			/codon_start=1
				/product="Fragment C-ter de la chimere SAH-Ukl-135"
				/protein_id="CA88669.1"
				/db_xref="GI:6088667"
				/translation="LGLSNELHQVPSMCDCLNGSTCYVSNKFTSIHMCNCPKRTGGOH
				CEIDNKRSTCTEGNCHFYRGKASTIDTKRCPCLPMSATVLOOTYHAHRSDALQQLSLGRH
				NYCNPNRNRPMQVYVGLAPLVQECMHVDCADGK"
BASE COUNT	117 a	109 c	107 g	90 t
ORIGIN				
Query Match	99.4%	Score 256.4;	DB 6;	Length 423;
Best Local Similarity	99.6%	Pred. NO. 1.7e-66;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
OY	1	AAACCTCTATGAGGGGAAATGGTCACTTTTACCAGAGAAAGGCCAGCACTGACACATG	60	
Db	153	AAACCTCTATGAGGGGAAATGGTCACTTTTACCAGAGAAAGGCCAGCACTGACACATG	212	
OY	61	GGCCGGCCCTGAGCTCCCTGGAACTCTGCACATGCTCTTCAGCAACGATACATGCCAC	120	
Db	213	GGCCGGCCCTGAGCTCCCTGGAACTCTGCACATGCTCTTCAGCAACGATACATGCCAC	272	
OY	121	AGATGTATGCTCTTCACCTGGAGCTGGGGGAAACATAATTTACTGCAGAACCCAGAAC	180	
Db	273	AGATGTATGCTCTTCACCTGGAGCTGGGGGAAACATAATTTACTGCAGAACCCAGAAC	332	
OY	181	CGGAGCGCACCTCGTGTGTATGTGAGAGTGGGCGCTTAAGCGCTGTCCAAAGATGCATG	240	
Db	333	CGGAGCGCACCTCGTGTGTATGTGAGAGTGGGCGCTTAAGCGCTGTCCAAAGATGCATG	392	
OY	241	GTGCATGACTGGCGCAT	258	
Db	393	GTGCATGACTGGCGCAT	410	
RESULT 2				
LOCUS	E02578	1236 bp	DNA	linear
DEFINITION	E02578			
ACCESSION	E02578.1	GI:2170808		
VERSION	JP 1990227075-A/6.			
KEYWORDS	synthetic construct.			
SOURCE	synthetic construct.			
ORGANISM	artificial sequences.			
REFERENCE	1 (bases 1 to 1236)			
AUTHORS	Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Ito,S.			
TITLE	NEW POLYPEPTIDE			
JOURNAL	Patent: JP 1990227075-A 6 10-SEP-1990;			
COMMENT	KYOMA HAKKO KOGYO CO LTD			
	OS Artificial gene			
	OC Artificial sequence; Genes.			
	PN JP 1990227075-A/6			
	PD 10-SEP-1990			
	PF 28-SEP-1989 JP 198925097			
	PR 29-SEP-1988 JP 88P 245705			
	PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEOSHI, PI			
	SATO MORIYUKI,			
	PI ITOU SEIGA			
	PC C12N5/72.C07K13/00.C07K15/14.C12N1/21.C12N5/10.C12N9/64, PC			
	C12N15/27			
	PC C12N15/58.C12N15/70.C12N15/85.C12P1/02.C12P21/02.C12N1/21,			
	PC C12R1/19)			
	PC (C12P21/02.C12R1/19).(C12P21/02.C12R1/91);			
	CC strandedness: Double;			
	CC topology: Linear;			
	CC hypothetical: No;			

FEATURES	Source	Location/Qualifiers
anti-sense: No;		
Key		
mat-peptide	1..1233	
FT		
FT	derivative(provK-S1) which is FT	resistant to
protease		
location/Qualifiers		
1..1236		
/organism="synthetic construct"		
/db_xref="taxon:32630"		
BASE COUNT	321 a 341 c 318 g 256 t	
ORIGIN		
Query Match	99.48;	Score 256.4; DB 6; Length 1236;
Best Local Similarity	99.68;	Fred. No. 1.6e-06;
Matches 257;	Conservative 0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	AAACCTCCTATGAGGGGAATGTCACCTTTTACCGAGGAAAGGCCAGCAGTACACCATG 60
Db	142	AAACCTCCTATGAGGGGAATGTCACCTTTTACCGAGGAAAGGCCAGCAGTACACCATG 201
QY	61	GGCCGCGCCCTGCGCTGCGCTGGAACCTGTCACCTGCTTCAGCAAAAGTACCATGCCAC 120
Db	202	GGCCGCGCCCTGCGCTGCGCTGGAACCTGTCACCTGCTTCAGCAAAAGTACCATGCCAC 261
QY	121	AGATCTAATGTCCTCTCAGCTGGGCTGGGGAACATAATTAATGTCAGGAAACCCAGACAAC 180
Db	262	AGATCTAATGTCCTCTCAGCTGGGCTGGGGAACATAATTAATGTCAGGAAACCCAGACAAC 321
QY	181	CGAGAGCCACCCCTGGTGTATGTGAGAGTGGGCGCTTAAGCGCTTGCCAGAGTGCATG 240
Db	322	CGAGAGCCACCCCTGGTGTATGTGAGAGTGGGCGCTTAAGCGCTTGCCAGAGTGCATG 381
QY	241	GTGCATGACTGGCGCAGAT 258
Db	382	GTGCATGACTGGCGCAGAT 399
RESULT 3		
LOCUS	E02579	1236 bp DNA linear PAT 29-SEP-1997
DEFINITION	Human prourokinase derivative.	
ACCESSION	E02579	
VERSION	E02579.1 GI:2170809	
KEYWORDS	JP 1990227075-A/7.	
SOURCE	synthetic construct.	
ORGANISM	synthetic construct.	
REFERENCE	artificial sequence.	
AUTHORS	1 (bases 1 to 1236)	
TITLE	Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Itou,S.	
JOURNAL	NEW POLYPEPTIDE	
COMMENT	Patent: JP 1990227075-A 7 10-SEP-1990;	
	KYOMA HAKKO KOGYO CO LTD	
	OS Artificial gene	
	OC Artificial sequence; Genes.	
	PN JP 1990227075-A/7	
	PD 10-SEP-1990	
	PF 28-SEP-1989 JP 1989253097	
	PR 29-SEP-1988 JP 88P 245705	
	PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI	
	SATO MORIYUKI,	
	PI ITOU SEIGA	
	PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N1/64, PC	
	C12N15/27	
	PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,C12N1/21,	
	PC C12R1:19)	
	PC C12P21/02,C12R1:19),(C12P21/02,C12R1:91);	
	CC strandedness: Double;	
	CC topology: linear;	
	CC hypothetical: No;	
	CC anti-sense: No;	
	key	Location/Qualifiers



DB	142	AAAACCTGCTATGAGGGGAAATGTCACCTTTACCGGAGAAAGGCCAGCAGTACACCATG	201
Qy	61	GGCGGGCCCTGCGTGGCCCTCCCTGGAACTCTGCGCACTGTCTTCAGCAAAACGTACCATGCCAC	120
Db	202	GGCGGGCCCTGCGTGGCCCTCCCTGGAACTCTGCGCACTGTCTTCAGCAAAACGTACCATGCCAC	261
Qy	121	AGATCTAATGCTCTTCAGCTGGGCGCTGGGGAAACAATAATTACTGCAAGAACCCAGACAAC	180
Db	262	AATCTTGATGCTCTTCAGCTGGGCGCTGGGGAAACAATAATTACTGCAAGAACCCAGACAAC	321
Qy	181	CGGAGCGCACCCCTGGTGCATATGTGCAGGTGGGGCCCTTAAAGCCGCTTGTCCAGAGTGCATG	240
Db	322	CGGAGCGCACCCCTGGTGCATATGTGCAGGTGGGGCCCTTAAAGCCGCTTGTCCAGAGTGCATG	381
Qy	241	GTGCATGACTGGCGCAGAT	258
Db	382	GTGCATGACTGGCGCAGAT	399
RESULT 6			
LOCUS	E02710		
DEFINITION	DNA sequence coding for prouokinase derivative, /K-63.		
ACCESSION	E02710		
VERSION	E02710.1		
KEYWORDS	JP 1991022979-A/3.		
SOURCE	Unidentified.		
ORGANISM	Hepatitis B virus		
REFERENCE	Viruses; Retrod viruses; Hepadnaviridae; Orthohepadnavirus.		
AUTHORS	1 (bases 1 to 1236)		
TITLE	Yasunura,S., Nishi,T. and Itou,S.		
JOURNAL	NOVEL PLASMINOGEN-ACTIVATION FACTOR		
COMMENT	Patent: JP 1991022979-A/3 31-JAN-1991;		
	KYOMA HAKKO KOGYO CO LTD		
	PN JP 1991022979-A/3		

```

PD      31-JAN-1991
PI      19-JUN-1989 JP 1989156302
PI      YASUDORA SHIGEOSHI, NISHI TATSUYA, ITOU SEIGA PC
C12M9/7.2,C12N15/58,(C12M9/7.2,C12R1:19));
CC      strandedness: Double;
CC      topology: linear;
CC      hypothetical: No;
CC      anti-sense: No;
CC      *source: library-plasmid pUKS3;
FH      Location/Qualifiers
FH      CDS
FT      1..1236
FT      /product='Prourokinase derivative,UK-S3' FT
modified_base 457..459
FT      /note='ctg' is replaced by 'aat' FT
modified_base 463..465
FT      /note='ccc' is replaced by 'acg' FT
FEATURES
source
Location/Qualifiers
1..1236
/organism='Hepatitis B virus'
/db_xref='taxon:10407'
BASE COUNT      322 a      338 c      317 g      259 t
ORIGIN
Query Match      99.4%; Score 256.4; DB 6; Length 1236;
Best Local Similarity 99.6%; Pred. No 1.6e-66;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 AAACCTCTATGAGGGGAATGCTACTTTTACGAGAGAAAGCCAGCACTGACCATG 60
|||||
Db 142 AAACCTCTATGAGGGGAATGCTACTTTTACGAGAGAAAGCCAGCACTGACCATG 201
61 GCGGGGCGCTGCGCTGCGCTGGAAGCTGCGCATGTCTTCAGCAAAAGCTACATGCCAC 120
|||||
Db 202 GCGGGGCGCTGCGCTGCGCTGGAAGCTGCGCATGTCTTCAGCAAAAGCTACATGCCAC 261
121 AGATCTAATGCTCTTCACTGGGCGCTGGGGAACAATATTAATCTACGAGAACCCAGACAC 180
|||||

```

Accession	Sequence	Length
Db	ACATCTGATGTCCTTTCACACTGGGCGCTGGGGAAACATAATATTCTGCAGAGAACCCAGACAAAC	322
Qy	181 CGAGAGCGCAACCTGGTGGTATATGTGCAGATGGGGCCCTAAAGCCGCTTTCACAAAGATGCATG	240
Db	322 CGAGAGCGCAACCTGGTGGTATATGTGCAGATGGGGCCCTAAAGCCGCTTTCACAAAGATGCATG	381
Qy	241 GTGCATGACTGGCGCAGAT	258
Db	382 GTGCATGACTGGCGCAGAT	399

RESULT 7			
E02832			
LOCUS	1236 bp	RNA	linear
DEFINITION	DNA encoding human pro-urokinase.		
ACCESSION	E02832		
VERSION	E02832.1		
KEYWORDS	GI:2171060		
SOURCE	JF 1991087180-A/1.		
	Homo sapiens.		

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Emmariyola; Metacoza; Chordata; Craniata; Vertebrata; Eucelostomata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1236)	Tanabe,T., Amatsuyu,Y., Kasai,S., Hirose,M., Morita,M., Kawabe,H. and Arimura,H.	MUTANT HUMAN-PRODRONINASE, ITS PRODUCTION, DNA SEQUENCE, PLASMID AND HOST	JOURNAL	
Patent: JP 1991087180-A 1 11-APR-1991;				
GREEN CROSS CORP:THE				
OS Homo sapiens (human)				

PN JP 1991087180-A/1  
PD 11-APR-1991  
PE 16-FEB-1990 JP 1990036809  
PR 18-MAY-1989 JP 89P 128434  
PI TANABE TOSHIZUMI, AMAISUJI YASUO, KASAI SHUNJI, HIROSE  
MASAKI, PI MORITA MASANORI, KAMABE HARUHIDE, ARIMURA HIROBUMI PC  
C12N9/72.C12N5/50.C12N5/58;

FEATURES	
CC	strandedness: Double;
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No;
FH	Key
FH	Location/Qualifiers
FT	CDS
FT	1. .1236
FT	/product='pro-urokinase'
FT	mat_peptide
FT	1. .1233
FT	/product='pro-urokinase'
FT	mutation
FT	/product='pro-urokinase mutant'
FT	replace(70. .72,'AAC')
FT	replace(70. .72,'TCC')
FT	/product='pro-urokinase mutant'
FT	replace(85. .87,'TCC')
FT	/product='pro-urokinase mutant'
FT	1. .147
FT	/note='epidermal growth factor domain'
FT	Location/Qualifiers
FT	Region

source	1. .1236	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT	318 a	342 c	319 g 257 t
ORIGIN			
Query Match		99.4%	Score 256.4; DB 6; Length 1236;
Best Local Similarity		99.66%;	Prod. No. 1.6e-66;
Matches 257; Conservative	0;	Mismatches 1;	Indels 0; Gaps
QY	1	AAACCTCTATGAGGGGAAATGTCACCTTTTACCGAGAAAGCCAGCAGCATGACACATG	60
Db	142	AAACCTCTATGAGGGGAAATGTCACCTTTTACCGAGAAAGCCAGCAGCATGACACATG	201
QY	61	GGCCGGCCCTCCAGCCCTGGAACCTGCGACAGTGTCTTCACACAAAGCTACCATGCCCCAC	120
Db	202	GGCCGGCCCTCCAGCCCTGGAACCTGCGACAGTGTCTTCACACAAAGCTACCATGCCCCAC	261

Db	202	GGCGGGGCGCCCGCCGCGCCGTGGAACTCTGACACGTCTCTTGACGAAAGTACATGCGCAC	261
Oy	121	AGATCTAATGCTCTTTCACTGCGGCGCTGGGAAACATATATCTCTCAGAAACCCAGACAAAC	180
Db	262	AGATCTGATGCTCTTTCACGTGGGCGCTGGGAAACATATATCTCTCAGAAACCCAGACAAAC	321
Oy	181	CGAGAGCCACCTCGTGTATGTCTCAGATGGGCGCTAAAGCGCGTTGTCTCAAGAGTGCATG	240
Db	322	CGAGAGCGACCTCGTGTATGTCTCAGATGGGCGCTAAAGCGCGTTGTCTCAAGAGTGCATG	381
Oy	241	GTGCATGACTGCGCAGAT	258

RESULT	9		
E02577			
LOCUS	E02577	1396 bp	RNA
DEFINITION	DNA encoding human prourokinase.		
ACCESSION	E02577		
VERSION	E02577.1 GI:2170807		
KEYWORDS	CP 1990227075-A/5.		
			PAT 29-SEP-1997

ORGANISM	REFERENCE
Homo sapiens	Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Eukariota	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Mammalia	1 (bases 1 to 1296)
1 (bases 1 to 1296)	Sasaki,K., Nishi,T., Yasunura,S., Seto,M. and Itou,S.
Sasaki,K., Nishi,T., Yasunura,S., Seto,M. and Itou,S.	NEW POLYPEPTIDE
NEW POLYPEPTIDE	Patent: JP 1990227075-A 5 10-SEP-1990;
Patent: JP 1990227075-A 5 10-SEP-1990;	KYOWA HAKKO KOGYO CO LTD

PN JP 1990327075-A/5  
PD 10-SEP-1990  
PF 28-SEP-1989 JP 1989253097  
PR 29-SEP-1988 JP 88P 245705  
PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, P  
MORIYUKI,  
SAITO  
ITOU SEIGA  
PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64,

PC (C12N1:5/58, C12N1:5/70, C12N1:5/85, C12P21/02, C12P22/02, (C12N1/21  
C12R1:19), C12R1:19), (C12P21/02, C12R1:91);  
PC (C12P21/02, C12R1:19);  
CC strandiness: Double;  
CC topological: linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC anti-sense: lincRNA-Derlet1+ 562 cell);

	*source: clone=DKAII;
CC	Location/Qualifiers
FH	Key
FH	
FT	
FT	
CDS	1..1296
FT	/product='human prourokinase precursor'
FT	.1293
mat_peptide	/product='human peoukoinase'.
FT	
FEATURES	Location/Qualifiers

BASE COUNT	327 a	361 c	337 g	271 t
ORIGIN	/Organism="Homo sapiens" /db_xref="taxon:9606"			
Query Match	99.4%	Score 256.4	DB 6	Length 1296
Best Local Similarity	99.6%	Pred. No. 1.6e-65		
Matches 257: Conservative	0	Mismatches 1	Indels 0	Gaps

QY	1	AAAACCTGCTATGAGGGGAAATGTCACCTTTTACGAGGAAAGCCAGCAGTACACACATG	60
Db	202	AAAACCTGCTATGAGGGGAAATGTCACCTTTTACGAGGAAAGCCAGCAGTACACACATG	26
QY	61	GGCGGGCCCTGCGCTGCGTGAACCTGCGCACTGTCCTTACAGCAAACTTACATGCGCAC	120

Db 262 GGGCGGGCCCTGCGCTGCGCTGGAACCTGCGACACGTCTTCAGCAAAAGCTACATGCCAC 321  
Qy 121 AGATCTAATGCTCTTCAGCTGGGCGGCAAAACATATTAATGACAGAAACCCAGACAC 180  
Db 322 AGATCTAATGCTCTTCAGCTGGGCGGCAAAACATATTAATGACAGAAACCCAGACAC 381  
Qy 181 CGGAGGCGACCCCTGCTGCTATGTCAGAGTGGGCGCTAAAGCCGCTTGTCAGAGTGCATG 240  
Db 382 CGGAGGCGACCCCTGCTGCTATGTCAGAGTGGGCGCTAAAGCCGCTTGTCAGAGTGCATG 441  
Qy 241 GTGCATGACTGGCGAGAT 258  
Db 442 GTGCATGACTGGCGAGAT 459  
RESULT 10  
LOCUS E02647 1296 bp RNA linear PAT 29-SEP-1997  
DEFINITION DNA sequence coding for pro-urokinase.  
ACCESSION E02647  
VERSION E02647.1 GI:2170875  
KEYWORDS JP 1990257891-A/2.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1296)  
AUTHORS Miyajima, H., Sasaki, K. and Itou, S.  
TITLE PRODUCTION OF PROTEIN BY RECOMBINANT ANIMAL CELL  
JOURNAL Patent: JP 1990257891-A 2 18-OCT-1990;  
KYOWA HAKKO KOGYO CO LTD  
OS Homo sapiens (human)  
PN JP 1990257891-A/2  
PD 18-OCT-1990  
PE 31-MAR-1989 JP 1989078573  
PI MIYAJI HIROMASA, SASAKI KATSUTOSHI, ITOU SEIGA PC  
C12N1/00, C12P21/02//C12N5/10, C12N15/09, (C12P21/00, C12R1.91), PC  
(C12N5/10,  
PC C12R1.91);  
CC strandness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: clone-pvkl;  
FH key Location/Qualifiers  
FT CDS 1..1296  
FEATURES  
source Location/Qualifiers  
1..1296  
/organism="Homo sapiens"  
BASE COUNT 327 a 361 c 337 g 271 t  
ORIGIN  
Query Match 99.4%; Score 256.4; DB 6; Length 1296;  
Best Local Similarity 99.6%; Pred. No. 1.6e-66;  
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 AAAACCTGCTATGAGGGGAATGTCACCTTTACCGAGAAAGCCAGACATGACCATG 60  
Db 202 AAAACCTGCTATGAGGGGAATGTCACCTTTACCGAGAAAGCCAGACATGACCATG 261  
Qy 61 GGGCGGCGCTGCTGCGCTGCGCAACTGCGCACTGCTTCAGCAAAAGCTACATGCCAC 120  
Db 262 GGGCGGCGCTGCTGCGCTGCGCAACTGCTTCAGCAAAAGCTACATGCCAC 321  
Qy 121 AGATCTAATGCTCTTCAGCTGGGCGGCAAAACATATTAATGACAGAAACCCAGACAC 180  
Db 322 AGATCTAATGCTCTTCAGCTGGGCGGCAAAACATATTAATGACAGAAACCCAGACAC 381  
Qy 181 CGGAGGCGACCCCTGCTGCTATGTCAGAGTGGGCGCTAAAGCCGCTTGTCAGAGTGCATG 240  
Db 442 GTGCATGACTGGCGAGAT 258  
RESULT 12  
LOCUS E02711 1296 bp DNA linear PAT 29-SEP-1997

Db 382 CGGAGGCGACCCCTGCTGCTATGTCAGAGTGGGCGCTAAAGCCGCTTGTCAGAGTGCATG 441  
Qy 241 GTGCATGACTGGCGAGAT 258  
Db 442 GTGCATGACTGGCGAGAT 459  
RESULT 11  
LOCUS E02649 1296 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA sequence coding for human pro-urokinase.  
ACCESSION E02649  
VERSION E02649.1 GI:2170877  
KEYWORDS JP 1990261386-A/1.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1296)  
AUTHORS Sekine, S., Itou, S. and Katsuki, M.  
TITLE RECOMBINANT VECTOR  
JOURNAL Patent: JP 1990261386-A 1 24-OCT-1990;  
KYOWA HAKKO KOGYO CO LTD, JITSUREN DOUBUTSU CHUO KENKYUSHO  
OS Homo sapiens (human)  
PN JP 1990261386-A/1  
PD 24-OCT-1990  
PE 31-MAR-1989 JP 1989078574  
PI SEKINE SUSUMU, ITOU SEIGA, KATSUKI MOTOYA  
PC C12N1/00, C12N15/85, A01K67/027, C12N1/21, C12N9/72//C12N15/58, (C12N1/21,  
PC C12R1.19);  
CC strandness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH key Location/Qualifiers  
FT CDS 1..1296  
FEATURES  
source Location/Qualifiers  
1..1296  
/organism="Homo sapiens"  
BASE COUNT 327 a 360 c 338 g 271 t  
ORIGIN  
Query Match 99.4%; Score 256.4; DB 6; Length 1296;  
Best Local Similarity 99.6%; Pred. No. 1.6e-66;  
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 AAAACCTGCTATGAGGGGAATGTCACCTTTACCGAGAAAGCCAGACATGACCATG 60  
Db 202 AAAACCTGCTATGAGGGGAATGTCACCTTTACCGAGAAAGCCAGACATGACCATG 261  
Qy 61 GGGCGGCGCTGCTGCGCTGCGCAACTGCGCACTGCTTCAGCAAAAGCTACATGCCAC 120  
Db 262 GGGCGGCGCTGCTGCGCTGCGCAACTGCGCACTGCTTCAGCAAAAGCTACATGCCAC 321  
Qy 121 AGATCTAATGCTCTTCAGCTGGGCGGCAAAACATATTAATGACAGAAACCCAGACAC 180  
Db 322 AGATCTAATGCTCTTCAGCTGGGCGGCAAAACATATTAATGACAGAAACCCAGACAC 381  
Qy 181 CGGAGGCGACCCCTGCTGCTATGTCAGAGTGGGCGCTAAAGCCGCTTGTCAGAGTGCATG 240  
Db 382 CGGAGGCGACCCCTGCTGCTATGTCAGAGTGGGCGCTAAAGCCGCTTGTCAGAGTGCATG 441  
Qy 241 GTGCATGACTGGCGAGAT 258  
Db 442 GTGCATGACTGGCGAGAT 459  
RESULT 12  
LOCUS E02711 1296 bp DNA linear PAT 29-SEP-1997

DEFINITION DNA sequence coding for native prourokinase.  
 ACCESSION E02711  
 VERSION E02711.1 GI:2170939  
 KEYWORDS JP 1991022979-A/4.  
 SOURCE Unidentified.  
 ORGANISM Hepatitis B virus  
 Viruses: Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 1 (bases 1 to 1296)  
 YASUMURA, S., NISHI, T., and ITOU, S.  
 NOVEL PLASMINOGEN-ACTIVATION FACTOR  
 Patent: JP 1991022979-A 4 31-JAN-1991;  
 KYOMA HAKKO KOGYO CO LTD  
 PN JP 1991022979-A/4  
 PD 31-JAN-1991  
 PF 19-JUN-1989 JP 1989156302  
 PI YASUMURA SHIGEYOSHI, NISHI TATSUYA, ITOU SEIGA PC  
 C12N9/72.C12N15/58.(C12N9/72.C12R1:19);  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 FH key Location/Qualifiers  
 FT sig\_peptide 1..60  
 FT /product='Native prourokinase signal peptide'  
 FT CDS 61..1296  
 FT /product='Native prourokinase' FT  
 FT mat\_peptide 61..192  
 FT /note='Growth factor domain'  
 FT mat\_peptide 193..459  
 FT /note='taringul domain'  
 FT mat\_peptide 460..1293  
 FT /note='Protease domain'.  
 FEATURES Location/Qualifiers  
 source 1..1296  
 /organism='Hepatitis B virus'  
 /db\_xref='taxon:10407'  
 BASE COUNT 327 a 361 c 337 g 271 t  
 ORIGIN  
 Query Match 99.4%; Score 256.4; DB 6; Length 1296;  
 Best Local Similarity 99.6%; Pred. No. 1.6e-66;  
 Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AAAACCTGCTATGAGGGGAATGTCATTTCACGAGAAAGCCAGACATGACCATG 60  
 DB 202 AAAACCTGCTATGAGGGGAATGTCATTTCACGAGAAAGCCAGACATGACCATG 261  
 QY 61 GGCCGGCCCTGCTGCTGAGAACTGTCACCTTCAGCAAAAGTACCATGCCCCAC 120  
 DB 262 GGCCGGCCCTGCTGCTGAGAACTGTCACCTTCAGCAAAAGTACCATGCCCCAC 321  
 QY 121 AGATCTAATGCTCTTACGCTGGGCTGGGAAACATTAATTCTGACGAAACCCAGACAAC 180  
 DB 322 AGATCTAATGCTCTTACGCTGGGCTGGGAAACATTAATTCTGACGAAACCCAGACAAC 381  
 QY 181 CGAGAGGCAACCTGCTGCTATGTCAGGTGGGCTTAAAGCCCTGTCACAAAGAGTCATG 240  
 DB 382 CGAGAGGCAACCTGCTGCTATGTCAGGTGGGCTTAAAGCCCTGTCACAAAGAGTCATG 441  
 QY 241 GTGCATGACTGCGCAGAT 258  
 DB 442 GTGCATGACTGCGCAGAT 459  
 RESULT 13  
 LOCUS E06064 1296 bp RNA linear PAT 29-SEP-1997  
 DEFINITION DNA encoding human urokinase.  
 ACCESSION E06064  
 VERSION E06064.1 GI:2174251  
 KEYWORDS JP 1993336965-A/2.  
 SOURCE Homo sapiens.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1296)  
 YASUMURA, S., YAMAMOTO, Y., HASEGAWA, M., HIGO, K., KUBO, K. and  
 KAWABARA, T.  
 HUMAN PROUROKINASE DERIVATIVE  
 Patent: JP 1993336965-A 2 21-DEC-1993;  
 KYOMA HAKKO KOGYO CO LTD  
 PN JP 1993336965-A/2  
 OS Homo sapiens (human)  
 PD 21-DEC-1993  
 PF 17-OCT-1991 JP 1991269615  
 PI YASUMURA SHIGEYOSHI, YAMAMOTO YOSHINORI, HASEGAWA MAMORU, PI  
 HIGO KATSUYA  
 PI KUBO KAZUHIRO, KAWABARA TAKASHI  
 PC C12N9/72.C12N1/21.C12N15/58;  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC \*source: strain-Detroit562;  
 CC \*source: clone-pUKL, pUKL1;  
 FH key Location/Qualifiers  
 FT CDS 1..1296  
 FT /product='human prourokinase' FT sig\_peptide  
 FT 1..60  
 FT mutation replace(88..195,'')  
 FT /note='UK-deltaGSI'  
 FT mutation replace(540..541,'aa')  
 FT /note='UK-deltaGSI'.  
 FEATURES Location/Qualifiers  
 source 1..1296  
 /organism='Homo sapiens'  
 /db\_xref='taxon:9606'  
 BASE COUNT 327 a 361 c 337 g 271 t  
 ORIGIN  
 Query Match 99.4%; Score 256.4; DB 6; Length 1296;  
 Best Local Similarity 99.6%; Pred. No. 1.6e-66;  
 Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AAAACCTGCTATGAGGGGAATGTCATTTCACGAGAAAGCCAGACATGACCATG 60  
 DB 202 AAAACCTGCTATGAGGGGAATGTCATTTCACGAGAAAGCCAGACATGACCATG 261  
 QY 61 GGCCGGCCCTGCTGCTGAGAACTGTCACCTTCAGCAAAAGTACCATGCCCCAC 120  
 DB 262 GGCCGGCCCTGCTGCTGAGAACTGTCACCTTCAGCAAAAGTACCATGCCCCAC 321  
 QY 121 AGATCTAATGCTCTTACGCTGGGCTGGGAAACATTAATTCTGACGAAACCCAGACAAC 180  
 DB 322 AGATCTAATGCTCTTACGCTGGGCTGGGAAACATTAATTCTGACGAAACCCAGACAAC 381  
 QY 181 CGAGAGGCAACCTGCTGCTATGTCAGGTGGGCTTAAAGCCCTGTCACAAAGAGTCATG 240  
 DB 382 CGAGAGGCAACCTGCTGCTATGTCAGGTGGGCTTAAAGCCCTGTCACAAAGAGTCATG 441  
 QY 241 GTGCATGACTGCGCAGAT 258  
 DB 442 GTGCATGACTGCGCAGAT 459  
 RESULT 14  
 LOCUS AX451990 1359 bp DNA linear PAT 06-JUL-2002  
 DEFINITION Sequence 3 from Patent WO0244393.  
 ACCESSION AX451990  
 VERSION AX451990.1 GI:21711991  
 KEYWORDS human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Tue Nov 12 18:06:49 2002

pct-us02-27855-2.rge

Page 8

REFERENCE	1	Armenadara, Borunda, J. and Aguilar Cordova, E.
AUTHORS		Recombinant viral and non-viral vectors containing the human
TITLE		urokinase plasminogen activator gene and its utilization in the
		treatment of various types of hepatic, renal, pulmonary, pancreatic
JOURNAL		and cardiac fibrosis and hypertrophic scars
		Patent: WO 024393-A 3 06-JUN-2002;
FEATURES		TGT LMB S A DE C V (MX)
SOURCE		Location/Qualifiers
		1. 1359
BASE COUNT	358 a	374 c 355 g 272 t
ORIGIN		
Query Match	99.4%	Score 256.4; DB 6; Length 1359;
Best Local Similarity	99.6%	Pred. No. 1.6e-66;
Matches 257; Conservative	0; Mismatches 1; Indels	0; Gaps 0;
QY	1	AAACCTGCTATGAGGGAGATGCTCTTTACCGAGAAAGCCAGCAGCTACACCATG 60
Db	228	AAAACTCTCTATGAGGGAGATGCTCTTTACCGAGAAAGCCAGCAGCTACACCATG 287
QY	61	GGCCGGCCCTGCGCTGCTGCAACTCTCCACTGCTCTTCAGCAACGTACCATGCCAC 120
Db	288	GGCCGGCCCTGCGCTGCTGCAACTCTCCACTGCTCTTCAGCAACGTACCATGCCAC 347
QY	121	AGACTTAATGCTCTTCACCTGGGCTGGGGAAACATTAATTACTGCAGAACCCACAGAC 180
Db	348	AGACTTGATGCTCTTCACCTGGGCTGGGGAAACATTAATTACTGCAGAACCCACAGAC 407
QY	181	CGAGAGCACCCTGGTGTATGTGAGAGTGGGCCCTAAAGCCGCTGTCACAGAGTCATG 240
Db	408	CGAGAGCACCCTGGTGTATGTGAGAGTGGGCCCTAAAGCCGCTGTCACAGAGTCATG 467
QY	241	GTGCATGACTGCGGAGAT 258
Db	468	GTGCATGACTGCGGAGAT 485
RESULT 15		
LOCUS	E00421	1394 bp RNA linear PAT 29-SEP-1997
DEFINITION	E00421	CDNA coding urokinase and its flanking region.
ACCESSION	E00421	GI:2168704
VERSION	E00421.1	GI:2168704
KEYWORDS	JP 1985180591-A/1.	
SOURCE		Homo sapiens.
ORGANISM		Homo sapiens
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
		1 (bases 1 to 1394)
REFERENCE		Hiramaetsu, T., Kaneda, T., Nagai, M., Ariuura, H., Nishida, M. and
AUTHORS		Suyama, T.
TITLE		DNA SEQUENCE, PLASMID AND HOST OF HUMAN UROKINASE
JOURNAL		Patent: JP 1985180591-A 1 14-SEP-1985;
COMMENT		GREEN CROSS CORP.:THE
		OS Homo sapiens (Human)
		PN JP 1985180591-A/1
		PD 14-SEP-1985
		PF 27-FEB-1984 JP 1984037119
		PI HIRAMAETSU TAKASHI, KANEDA TERUO, NAGAI MASANORI, PI
		HIROMITSU
		PI NISHIDA MASAYUKI, SUYAMA TADAKAZU
		PC C12M15/00, C07H21/04, C12M1/00//C12M9/72, (C12M1/00, C12M1/19); CC
		strandedness: Double;
		CC topology: Linear;
		CC hypothetical: No;
		CC anti-sense: No;
		CC *source: tissue_type=kidney;
		CC Feature is identified by similarity;
		PH Key Location/Qualifiers
		PH

FEATURES	source	location/Qualifiers
FT	sig_peptide	1..60
FT	mat_peptide	61..1296
FT	/Product='urokinase'	1..1296.
FT	CDS	1..1394
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
BASE COUNT	349 a 385 c 367 g 293 t	
ORIGIN		
Query Match	99.4%;	Score 256.4; DB: 6; Length 1394;
Best Local Similarity	99.6%;	Pred. No. 1.6e-66;
Matches 257; Conservative	0; Mismatches 1;	Indels 0; Gaps 0
QY	1 AAAACCTGCTATGAGGGGAATGTCACCTTTACGAGGAAAGCCGACGTGACACCATG 60	
Db	202 AAAACCTGCTATGAGGGGAATGTCACCTTTACGAGGAAAGCCGACGTGACACCATG 261	
QY	61 GGGCGGCGCCCTGCGCCCTGGAACTCTGGCACTGTCTTCAGCAAAAGTACCATGCCCAC 120	
Db	262 GGGCGGCGCCCTGCGCCCTGGAACTCTGGCACTGTCTTCAGCAAAAGTACCATGCCCAC 321	
QY	121 AGATCTAATGCTCTTACGTGAGCGCGCGGGGAACAATAATTACTCAGAGAACCCAGACAAC 180	
Db	322 AGATCTAATGCTCTTACGTGAGCGCGCGGGGAACAATAATTACTCAGAGAACCCAGACAAC 381	
QY	181 CGGAGGGGACCCGTGGGTGATGTGCGAGGGGGCCCTAAAGCCGTTGTCCAAAGATGCATG 240	
Db	382 CGGAGGGGACCCGTGGGTGATGTGCGAGGGGGCCCTAAAGCCGTTGTCCAAAGATGCATG 441	
QY	241 GTGCATGACTGGCGCAGAT 258	
Db	442 GTGCATGACTGGCGCAGAT 459	

Search completed: November 11, 2002, 12:24:38  
 Job time : 2627 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2002, 11:34:22 : Search time 231 Seconds  
(without alignments)  
2515.219 Million cell updates/sec

Title: PCT-US02-27855-2

Perfect score: 258  
Sequence: 1 aaacctgcatactaggggaa.....tgcgcatactgcgcagat 258

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.GeneSeq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.4	99.4	264	AAD27075	Human urokinase-ty
2	256.4	99.4	288	AAD27083	Human uPA kringle
3	256.4	99.4	405	AAD27078	Human urokinase-ty
4	256.4	99.4	429	AAD27082	Human uPA amino te
5	256.4	99.4	1137	11 AAT61672	CDNA encoding huma
6	256.4	99.4	1170	11 AAT61673	Human prourokinase
7	256.4	99.4	1206	11 AAT61674	Human prourokinase
8	256.4	99.4	1212	24 AAD27080	Human urokinase-ty
9	256.4	99.4	1236	11 AAT61671	Human native prour

10	256.4	99.4	1236	11 AA004486	Plasmid pUK33 enco
11	256.4	99.4	1236	11 AA006133	Sequence encoding
12	256.4	99.4	1236	11 AA006134	Sequence encoding
13	256.4	99.4	1236	11 AA006135	Sequence encoding
14	256.4	99.4	1236	11 AA010169	Encodes Pro-urokin
15	256.4	99.4	1236	12 AA010170	Encodes Pro-urokin
16	256.4	99.4	1236	24 AA027077	Human urokinase-ty
17	256.4	99.4	1236	24 AA027077	Human urokinase-ty
18	256.4	99.4	1296	11 AA006049	Plasid pUK1 pro-u
19	256.4	99.4	1296	15 AA055772	Pro-urokinase deri
20	256.4	99.4	1296	24 ABR66588	Human plasminogen
21	256.4	99.4	1473	10 AAN92037	Sequence of varian
22	256.4	99.4	1473	6 AAN50138	Sequence of the s1
23	256.4	99.4	1475	9 AAN50141	Sequence encoding
24	256.4	99.4	1475	9 AAN1558	Pro-UK structural
25	256.4	99.4	1475	20 AA224619	Human lung tumor a
26	256.4	99.4	1475	21 AAC65858	Human lung cancer-
27	256.4	99.4	1475	22 ABU49077	Human lung tumor
28	256.4	99.4	1964	22 AAH28220	Nucleotide sequenc
29	256.4	99.4	2266	24 AAD27855	Human uPA CDNA. H
30	256.4	99.4	2281	20 AAN93079	Sequence encoding
31	256.4	99.4	2294	20 AA224620	Human lung tumor a
32	256.4	99.4	2294	21 AAC65859	Human lung cancer-
33	256.4	99.4	2298	24 ABU49078	Human lung tumor
34	256.4	99.4	2299	7 AAN80981	Sequence of the s1
35	256.4	99.4	2301	8 AAN70390	Sequence encoding
36	256.4	99.4	2301	10 AAN91075	CDNA encoding huma
37	256.4	99.4	2303	10 AAN91740	DNA encoding natur
38	256.4	99.4	2303	15 AAN91740	Sequence of prouro
39	256.4	99.4	2304	4 AAN30030	Full length human
40	256.4	99.4	2304	8 AAN71338	Sequence encoding
41	256.4	99.4	2304	8 AAN71338	Modified prourokin
42	256.4	99.4	2304	8 AAN71368	Modified prourokin
43	256.4	99.4	2304	8 AAN71368	Modified prourokin
44	256.4	99.4	2377	13 ABU59583	Human pro-urokinas
45	256.4	99.4	2427	11 AAQ20360	Human pro-urokinas
					Human pro-urokinas

#### ALIGNMENTS

RESULT 1  
AAD27075  
ID AAD27075 standard; DNA: 264 BP.

XX  
AC AAD27075:  
XX  
DT 09-APR-2002 (first entry)  
XX

Human urokinase-type plasminogen activator (uPA) kringle DNA.

XX  
KW Human: urokinase-type plasminogen activator; uPA; therapy: hypertension;  
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
KW microvascular occlusion; angiosclerotic disorder; pulmonary fibrosis; asthma;  
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
KW clotting disorder; uterine contraction disorder; respiratory disease;  
KW male impotence; adult respiratory distress syndrome; ds.  
XX  
OS Homo sapiens.

XX  
FH Location/Qualifiers  
FH Key 1..264  
FH CDS /tag- a  
FT /product- "Human uPA kringle"  
FT /note- "CDS does not include start and stop codon"  
FT /partial

WO200197752-A2.

27-DEC-2001.

13-JUN-2001; 2001WO-US18976.



	RESULT 3
ID	AAD27078
XX	MAD27078 standard; DNA; 405 BP.
AC	
XX	AAD27078;
DT	09-APR-2002 (first entry)
XX	
DE	Human urokinase-type plasminogen activator amino terminal fragment DNA.
XX	
KW	Human: urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder; KM microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma; tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing; KW clotting disorder; uterine contraction disorder; respiratory disease; adult respiratory distress syndrome; amino terminal fragment; ATP; male impotence; ds. KN  OS Homo sapiens. XX XX FH Key Location/Qualifiers FT CDS 1..405 FT /tag= "a" FT /product= "Human uPA ATP" FT /note= "CDS does not include start and stop codon" FT /partial PN MO20019752-A2. XX PD 27-DEC-2001. XX PF 13-JUN-2001; 2001MO-US18976. XX PR 20-JUN-2000; 2000US-212874P. XX PA <unpe> UNIV PENNSYLVANIA. PI Clnes DB, Hlgazi AA: PI DR WPI: 2002-122240/J16. DR P-PDB: AAEI6545. XX PT Composition for modulating muscle cell and tissue contractility for treating atherosclerosis, asthma, hypertension, glaucoma, impotence, PT comprising domains from urokinase-type plasminogen activator - XX PS Claim 29; Fig 1M; 117pp: English. XX CC The invention relates to a composition comprising one or more domains of urokinase-type plasminogen activator (uPA). The composition is used to modulate the contractility and angiogenic activity of a mammalian muscle, endothelial cell or tissue. The composition is used for treating stroke, hypertension, hypotension, atherosclerosis, heart attack, microvascular occlusions, thrombotic microangiopathies, surgically induced thrombotic disorders, angioecnic disorders, pulmonary fibrosis, asthma, tumour cell invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma, diabetic retinopathy, wound healing, clotting disorder, uterine contraction disorder, male impotence, respiratory disease or condition such as asthma, adult respiratory distress syndrome, primary pulmonary hypertension, microvascular thrombotic occlusion, and a disorder associated with chronic intrapulmonary fibrin formation. The present sequence is human urokinase-type plasminogen activator (uPA) amino terminal fragment (ATF) DNA. XX SQ Sequence 405 BP; 113 A; 105 C; 104 G; 83 T; 0 other;
Query Match	99.4%; Score 256.4; DB 24; Length 405;
Best Local Similarity	99.6%; Pred.No.1,3e-34;
Matches 257; Conservative	0; Mismatches 1; Indels 0; Gaps 0
1 AAAAAGCTGCTATTAGGGGAATGTGTCCTTTTACCGAGAAGAAAGGCCAGACTGCACCACATG 60	

[illegible]



---



```

FH Key Location/Qualifiers
FT CDS 1..1236
FT /*tag= a
FT /product= prourokinase
FT /transl_except= pos:528..530, aa:Thr
FT /note= "the EGF domain is encoded by nucleotides
FT 28-126; novel variants of hPK which lack
FT at least part of the EGF domain are claimed"
FT Misc-difference 177
FT /note= "corresponds to TAC codon"
XX
XX EP398361-A.
XX
XX 22-NOV-1990.
XX
XX 18-MAY-1990; 90EP-0109472.
XX
XX 22-FEB-1990; 90JP-0042020.
XX
XX 18-MAY-1989; 89JP-0126433.
XX
XX 03-JUL-1986; 86JP-0156936.
XX
XX 18-FEB-1987; 87JP-0036495.
XX
XX 18-MAY-1989; 89JP-0126434.
XX
XX (GREC ) GREEN CROSS CORP.
XX
XX Alimura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
XX Morita M, Tanabe T;
XX
XX WPI: 1990-350146/47.
XX
XX P-PSDB; AAW13634.
XX
XX Human pro-urokinase variants - deficient in loop regions of
XX epidermal growth factor, showing long blood half-life, as
XX fibrinolytic agent
XX
XX Disclosure; Fig 1; 22pp; English.
XX
XX PS
XX
XX CC New variants of human prourokinase (hPK) comprise a hPK deficient
XX in (1) at least part of the first loop region of the epidermal growth
XX factor (EGF) domain; (11) at least part of the first loop and at
XX least part of the second loop; or (111) at least part of the third
XX loop. The hPK variants show an increased blood half-life comparable
XX to that of the whole EGF domain-deficient hPK variant and urokinase
XX while retaining the same properties as those of hPK. They have
XX potent thrombolytic activity and very little tendency to cause
XX spontaneous bleeding. The present sequence encodes the wild-type
XX hPK protein, including the EGF domain.
XX
XX SQ Sequence 1236 BP; 318 A; 341 C; 320 G; 257 T; 0 other;
XX
XX Query Match 99.4%; Score 256.4; DB 11; Length 1236;
XX Best Local Similarity 99.6%; Pred. No. 9.2e-35;
XX Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 AAAAAGCTGATAGAGGGAATGTCACCTTTACCGAGAAAGCCAGCATGACCATG 60
XX 142 AAAAAGCTGATAGAGGGAATGTCACCTTTACCGAGAAAGCCAGCATGACCATG 201
XX
XX 61 GGGCGGCGCTGCTGCGCTGGAACCTGTCACCTTCAGCAAAAGTACATGCCAC 120
XX 202 GGGCGGCGCTGCTGCGCTGGAACCTGTCACCTTCAGCAAAAGTACATGCCAC 261
XX
XX 121 AGATCTAATGCTCTTACGCTGGGCGCTGGGAAACATAATTACGCAAGAACCA 180
XX 262 AGATCTAATGCTCTTACGCTGGGCGCTGGGAAACATAATTACGCAAGAACCA 321
XX
XX 181 CGAGAGGCAAGCTGCTGCTGATGCAAGTGGGCTAAAGCGCTTCCAAAGTGCATG 240
XX 322 CGAGAGGCAAGCTGCTGCTGATGCAAGTGGGCTAAAGCGCTTCCAAAGTGCATG 381
XX
XX 241 GTGCATGACTGCGCAGAT 258
XX 382 GTGCATGACTGCGCAGAT 399
XX
XX Db

```

```

RESULT 10
AA004486
ID AA004486 standard; DNA: 1236 BP.
XX
XX AA004486;
XX
XX 04-OCT-1990 (first entry)
XX
XX Plasmid pUKS3 encoding UK-S3.
XX
XX Urokinase; glycosylation; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX
XX misc_difference 457..459
XX
XX /*tag= a
XX /label=synthetic-mutation
XX /note="old seq (ctg)"
XX
XX misc_difference 463..465
XX
XX /*tag= b
XX /label=synthetic-mutation
XX /note="old seq (ccc)"
XX
XX EP370205-A.
XX
XX 30-MAY-1990.
XX
XX 28-SEP-1989; 89EP-0117981.
XX
XX 29-SEP-1989; 89JP-0245705.
XX
XX (KYOW ) KYOWA HAKKA KK.
XX
XX Sasaki K, Nishi T, Yasumuru S, Sato M, Itoh S;
XX
XX WPI: 1990-165029/22.
XX
XX P-PSDB; AAR05117.
XX
XX Polypeptide(s) with added carbohydrate chains - formed by
XX modification of amino acid sequence, used to improve
XX physicochemical properties and/or activities.
XX
XX Disclosure; 30pp; English.
XX
XX PS
XX
XX CC The sequence encodes a deriv. of mature urokinase, designated
XX UK-S3 which has 2 amino acid substns. which result in an N-linked
XX glycosylation site giving the new protein improved stability and
XX activity.
XX
XX See also AA004481-85.
XX
XX SQ Sequence 1236 BP; 325 A; 336 C; 316 G; 259 T; 0 other;
XX
XX Query Match 99.4%; Score 256.4; DB 11; Length 1236;
XX Best Local Similarity 99.6%; Pred. No. 9.2e-35;
XX Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 AAAAAGCTGATAGAGGGAATGTCACCTTTACCGAGAAAGCCAGCATGACCATG 60
XX 142 AAAAAGCTGATAGAGGGAATGTCACCTTTACCGAGAAAGCCAGCATGACCATG 201
XX
XX 61 GGGCGGCGCTGCTGCGCTGGAACCTGTCACCTTCAGCAAAAGTACATGCCAC 120
XX 202 GGGCGGCGCTGCTGCGCTGGAACCTGTCACCTTCAGCAAAAGTACATGCCAC 261
XX
XX 121 AGATCTAATGCTCTTACGCTGGGCGCTGGGAAACATAATTACGCAAGAACCA 180
XX 262 AGATCTAATGCTCTTACGCTGGGCGCTGGGAAACATAATTACGCAAGAACCA 321
XX
XX 181 CGAGAGGCAAGCTGCTGCTGATGCAAGTGGGCTAAAGCGCTTCCAAAGTGCATG 240
XX
XX Db

```







**Tue Nov 12 18:06:51 2002**

pct-us02-27855-2.rng

Page 10

Accession	Sequence	Length
Db	GGCCGGCCCTCCCTCGCCCTGGAACTCTGCCACTGTCTCTTAGCAAAACGTACCATGCCAC	261
Qy	AGATCTAATGTCTCTTCAGCTGGGGCCCTGGGGAACATATTAATCTGACGAACCCAGACAC	180
Db	AGATCTGTATGCTCTTCACGTGGGCTTGGGGAACATATTAATCTGACGAACCCAGACAC	321
Qy	CGAGAGCGACCTCTGTGCTATGTGTCAGCTGAGCCCTAAAGCGCTGTGTCCAAAGTGCATG	240
Db	CGAGAGCGACCTCTGTGCTATGTGTCAGCTGAGCCCTAAAGCGCTGTGTCCAAAGTGCATG	381
Qy	GTGCATGACTCGCAGAT	258
Db	GTGCATGACTCGCAGAT	399

RESULT 15	
AAQ10170	
ID	AAQ10170 standard; DNA; 1236 BP.

AC	AAQ10170;
XX	
DT	18-MAR-1991 (first entry)

DE Encodes Pro-urokinase derivative UK-S3 with Asn(153) and Thr(155).  
XX  
XX pro-urokinase; UK-S3; plasminogen activator; myocardial infarction.  
KM cerebral thrombosis; ss.

Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	1.1236
FT		/*tag= a
FT		/product= UK-S3

PN EP405285-A.

02-JAN-1991.  
PD

AA 18-JUN-1990; 90EP-0111471.  
PF

AA 19-JUN-1989: 89JP-0156302.

XX  
PA ( KYOW ) KYOWA HAKKO KOGYO KK.

PI Yasamura S. Nishi T. Tto S:  
XX

XX WPB: 1991-008678/02

DR P-PSDB; AAR10058.  
XX

PT New plasminogen activator almost identical to natural  
pro-urokinase, is thrombolytic and used for

PT prophylaxis, treatment of cerebral thrombosis or myocardial

2  
L  
J  
J  
J  
J  
J

The *uvrS3* clone example of a plasmidogen activator which differs from natural human pro-urokinase at positions 153 and 155. There are corresponding codon changes at position 457-459 (Ileu substituted by CC Aaa) and at position 463 to 465 (Pro substituted by Thr). Of this DNA sequence, relative to the wild-type coding region, pBspA<sub>2</sub>, constructed from pUKL1, is cleaved with EcoRI and HindIII and a 3,4kb DNA fragment is obtained. pUKS1 is cleaved with HindIII and CfrI to give a 0,75kb fragment. The two fragments, together with two phosphorylated synthetic DNAs for site-specific mutation, were ligated together in the presence of T4 ligase to give pUK169 which has the sequence given here. See also AAQ010168 and AAQ010169.

Sequence 1236 BP; 322 A; 338 C; 317 G; 259 T; 0 other;

Query Match	99.48;	Score 256.4;	DB 12;	Length 1236;
Best Local Similarity	99.68;	Pred. No. 9.2e-35;		

Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	AAAACTGCTATGAGGGGAAATGTCACATTTTACCGAGAAAGGCCAGCACTACACCATG	60
QY	142	AAAACCTGCTATGAGGGGAAATGTCACATTTTACCGAGAAAGGCCAGCACTACACCATG	201
Db	61	GGCGGGGCGCTGGCTCCCGTGGAAACATCTGCGACATGTCCTTCACGAAACGATACATGCGCAC	120
QY	202	GCGCGGCGCTGGCTCCCGTGGAAACATCTGCGACATGTCCTTCACGAAACGATACATGCGCAC	261
Db	121	AGATCTTAATGTCCTTCACTAGGCGCTGGGGAAACATTAATTAATTCGACGAAACCGACACAC	180
QY	262	AGATCTTAATGTCCTTCACTAGGCGCTGGGGAAACATTAATTAATTAATTCGACGAAACCGACACAC	321
Db	181	CGAGAGCGAACCCTGTGCTATGTGCAAGGTGGGCCCTAAAGCCGCTTGTCCAGAGTGCATG	240
QY	322	CGAGAGCGAACCCTGTGCTATGTGCAAGGTGGGCCCTAAAGCCGCTTGTCCAGAGTGCATG	381
Db	241	GTCGATGACTGCGCAGAT	258
QY	382	GTCGATGACTGCGCAGAT	399

Search completed: November 11, 2002, 11:40:39  
Job time : 234 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2002, 11:36:36 : Search time 44 Seconds  
(without alignments)  
2080.333 Million cell updates/sec

Title: PCT-US02-27855-2

Perfect score: 258

Sequence: 1 aaacactgctatgaggggaa.....tggtcactgactgcagat 258

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCITS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.4	99.4	264	10	US-09-880-503-10
2	256.4	99.4	288	10	US-09-880-503-18
3	256.4	99.4	405	10	US-09-880-503-13
4	256.4	99.4	423	10	US-09-884-186-11
5	256.4	99.4	429	10	US-09-880-503-17
6	256.4	99.4	1212	10	US-09-880-503-15
7	256.4	99.4	1236	10	US-09-880-503-12
8	256.4	99.4	1475	10	US-09-735-705-122
9	256.4	99.4	1475	10	US-09-850-716A-122
10	256.4	99.4	1475	10	US-09-897-778-122
11	256.4	99.4	2294	10	US-09-735-705-123
12	256.4	99.4	2294	10	US-09-850-716A-123
13	256.4	99.4	2294	10	US-09-897-778-123
14	256.4	99.4	2036	10	US-09-954-456-552
15	256.4	99.4	2036	10	US-09-880-107-1612
16	256.4	99.4	1689	10	US-09-969-271-6
17	256.4	99.4	2519	10	US-09-969-271-5
18	256.4	99.4	2641	9	US-09-974-298-144
19	256.4	99.4	1683	10	US-09-912-559-1

20	55.2	21.4	1683	10	US-09-912-559-2	Sequence 2, Appl1
21	55.2	21.4	3008	10	US-09-880-107-1668	Sequence 1668, Ap
22	50.2	19.5	20556	10	US-09-880-107-3945	Sequence 3945, Ap
23	36.6	14.2	2893	10	US-09-848-288-2	Sequence 2, Appl1
24	33.2	12.9	4139	10	US-09-880-107-2279	Sequence 2279, Ap
25	32.4	12.6	2559	9	US-09-897-427A-5	Sequence 5, Appl1
26	32.4	12.6	2559	9	US-09-897-427A-7	Sequence 7, Appl1
27	32.4	12.6	2559	10	US-09-796-338A-15	Sequence 15, Appl
28	32.4	12.6	3489	10	US-09-796-338A-13	Sequence 13, Appl
29	32.2	12.5	9980	10	US-09-880-107-3352	Sequence 3352, Ap
30	31	12.0	3357	10	US-09-764-847-1478	Sequence 1478, Ap
31	31	12.0	3908	10	US-09-764-847-1477	Sequence 117, App
32	30.6	11.9	2372	10	US-09-925-302-319	Sequence 319, App
33	29.8	11.6	1343	10	US-09-822-849A-127	Sequence 127, App
34	29.8	11.6	1622	10	US-09-764-864-306	Sequence 306, App
35	29.8	11.6	1646	10	US-09-764-864-726	Sequence 726, App
36	29.6	11.5	431	10	US-09-864-761-13941	Sequence 13941, A
37	29.6	11.5	4237	10	US-09-962-832-117	Sequence 117, App
38	29.4	11.4	447	10	US-09-764-847-277	Sequence 277, App
39	29.4	11.4	2429	10	US-09-764-864-69	Sequence 69, Appl
40	29.4	11.4	10758	10	US-09-764-869-1647	Sequence 1647, Ap
41	29.4	11.4	10758	10	US-09-764-869-1647	Sequence 1650, Ap
42	29.4	11.4	10759	10	US-09-764-869-1649	Sequence 1649, Ap
43	29.2	11.3	200	10	US-09-864-761-24994	Sequence 24994, A
44	29.2	11.3	513	10	US-09-864-761-8260	Sequence 8260, Ap
45	29.2	11.3	3708	10	US-09-904-065-5	Sequence 5, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-880-503-10  
Sequence 10, Application US/09880503  
Patent No. US20020131964A1  
GENERAL INFORMATION:  
APPLICANT: CINES, Douglas B  
APPLICANT: HICAZI, Abd Al-Roof  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
TITLE OF INVENTION: TISSUE CONTRACTILITY  
FILE REFERENCE: 9596-331  
CURRENT APPLICATION NUMBER: US/09/880,503  
CURRENT FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/212,847  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 10  
LENGTH: 264  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-880-503-10

Query Match 99.4%; Score 256.4; DB 10; Length 264;  
Best Local Similarity 99.6%; Pred. No. 4e-77;  
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AAAACCTCTGTGAGGGAAATGCTCTTTTACCGAGGAAGCCACACTGACCATG 60  
DB 1 AAAACCTCTGTGAGGGAAATGCTCTTTTACCGAGGAAGCCACACTGACCATG 60  
QY 61 GCGCGCCCTCTGCGCCCTGAGACCTGCGCTCTTTCGCAAAAGTACATGCCAC 120  
DB 61 GCGCGCCCTCTGCGCCCTGAGACCTGCGCTCTTTCGCAAAAGTACATGCCAC 120  
QY 121 AGATCTAATGCTCTTACCTGAGGCTGGGGAACATATTACTGACAGAACCCAGACAC 180  
DB 121 AGATCTAATGCTCTTACCTGAGGCTGGGGAACATATTACTGACAGAACCCAGACAC 180  
QY 181 CGGAGGCGACCTGTGTGTATGTCAGGTGGGCTTAAAGCGCTTGTCCAAAGATGATG 240  
DB 181 CGGAGGCGACCTGTGTGTATGTCAGGTGGGCTTAAAGCGCTTGTCCAAAGATGATG 240

Tue Nov 12 18:06:52 2002

pct-us02-27855-2.rnpb

Page 2

OY	241	GTGCATGACTGGCGCAGAT	258
Db	241	GTGCATGACTGGCGCAGAT	258

RESULT 2  
US-09-880-503-18

```

: APPLICANT: CINES, Douglas B
: APPLICANT: HIGAZI, Abd Al-Roof
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
: TITLE OF INVENTION: TISSUE CONTRACTABILITY
: FILE REFERENCE: 9596-331
: CURRENT APPLICATION NUMBER: US/09/880,503
: CURRENT FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/212,847
: PRIOR FILING DATE: 2000-06-20
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 18
: LENGTH: 288
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-880-503-18

```

Query Match	99.48;	Score 256.4;	DB 10;	Length 288;
Best Local Similarity	99.68;	Pred No. 4.2e-77;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

Oy	1	AAACCTCTATGAGGGAAATGTCACCTTTACCGAGAAAGGCACACATGCATC	60
Db	1	AAACCTCTATGAGGGAAATGTCACCTTTACCGAGAAAGGCACACATGCATC	60
Oy	61	GGCGGGCCCTGACGCCCCGTGGAACTGTGCGACATGCTCCTTAGCAAAAGTACCATGCCAC	120
Db	61	GGCGGGCCCTGACGCCCCGTGGAACTGTGCGACATGCTCCTTAGCAAAAGTACCATGCCAC	120
Oy	121	AGATCTATGCTCTTTCACGCTGGGCGTGGGAAATATATATCTGACAGAAACCCAGAAAC	180
Db	121	AGATCTATGCTCTTTCACGCTGGGCGTGGGAAATATATATCTGACAGAAACCCAGAAAC	180
Oy	181	CGGAGGCGACCTCGTGTATGTACAGTGGGCGCTAAAGCCGCTGTTCACAAAGTGCATG	240
Db	181	CGGAGGCGACCTCGTGTATGTGTACAGTGGGCGCTAAAGCCGCTGTTCACAAAGTGCATG	240
Oy	241	GTCGATGACTGCGCAGAT	258
Db	241	GTCGATGACTGCGCAGAT	258

```

RESULT 3
US-09-880-503-13
; Sequence 13, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; TITLE OF INVENTION: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AN
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-880-503-13

```

Query Match	99.48;	Score 256.4;	DB 10;	Length 405;
Best Local Similarity	99.68;	Pred. No. 4.8e-77;		
Matches 257;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAAACTCTGTTAGAGGGAAATGTCACCTTTTACCGAGAAAAGGCACACTGACACCATG	60
Db	142	AAAACCTCTGTTAGAGGGAAATGTCACCTTTTACCGAGAAAAGGCACACTGACACCATG	201
QY	61	GGCGGGCCCTCAGCGCCCTGGAACTCTGCCACTCTCCTTGACGAAAGGTACCATGCCAC	120
Db	202	GGCGGGCCCTCAGCGCCCTGGAACTCTGCCACTCTCCTTGACGAAAGGTACCATGCCAC	261
QY	121	AGATCTATGCTCTTTCAGAGCTGGGCTGGGGAACAATATTAATCTCAGAGAACCCAGAAC	180
Db	262	AGATCTATGCTCTTTCAGAGCTGGGCTGGGGAACAATATTAATCTCAGAGAACCCAGAAC	321
QY	181	CGAGAGGCAACCTCTGCTATGTTCTCAGATGGGCGCTAAAGCGCTTTCAGAAAGTGCATG	240
Db	322	CGAGAGGCAACCTCTGCTATGTTCTCAGATGGGCGCTAAAGCGCTTTCAGAAAGTGCATG	381
QY	241	GTGCATGACTGCGCAGAT	258
Db	382	GTGCATGACTGCGCAGAT	399

RESULT 4  
US-09-984-186-11  
; Sequence 11, Application US/09984186  
; Patent No. US2002015101A1

APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES.

CONTAINING SAID POLY

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rotex Inc

STREET: 500 Arcoia Road, 3C43

CITY: Collegeville

STATE: PA

COUNTRY: USA

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: Macintosh
4  OPERATING SYSTEM: System 7.1
5  SOFTWARE: Word 5.1 (Patentin)
6
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/09/984,186
9  FILING DATE: 29-Oct-2001
10 CLASSIFICATION: <Unknown>

```

7 PRIOR APPLICATION DATA:  
 7 APPLICATION NUMBER: US/08/797,689  
 7 FILING DATE: 31-JAN-1997  
 7 APPLICATION NUMBER: US 08/256,927  
 7 FILING DATE: 28-JUL-1994  
 7 APPLICATION NUMBER: FR 92/01064  
 7 FILING DATE: 31-JAN-1992  
 7 APPLICATION NUMBER: PCT/FR93/00085  
 7 FILING DATE: 28-JAN-1993  
 7  
 7 ATTORNEY/AGENT INFORMATION:  
 7 NAME: SMITH Ph.D., Julie K  
 7 REGISTRATION NUMBER: P-38,619  
 7 REFERENCE/DOCKET NUMBER: S792006-US  
 7  
 7 TELECOMMUNICATION INFORMATION:  
 7 TELEPHONE: (610) 454-3839  
 7 TELEFAX: (610) 454-8808  
 7  
 7 INFORMATION FOR SEQ ID NO: 11:





Patent No. US20020147143A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong

Tue Nov 12 18:06:52 2002

pct-us02-27855-2.rmpb

Page 6

```

; APPLICANT: Marnierakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455c16
; CURRENT APPLICATION NUMBER: US/09/897,778
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for windows version 4.0
; SEQ ID NO: 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-123

Query Match
Best Local Similarity 99.4%; Score 256.4; DB 10; Length 2294;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 AAAACCTGATGAGGGGATGGTCACTTTACCGAGAAAGCCGACACTGACACCATG 60
DB 268 AAAACCTGATGAGGGGATGGTCACTTTACCGAGAAAGCCGACACTGACACCATG 327
QY 61 GGGCGGCGCTGCTGCTGGAACCTGCGCACTGCTTCCAGCAACGTAACATGCCAC 120
DB 328 GGGCGGCGCTGCTGCTGGAACCTGCGCACTGCTTCCAGCAACGTAACATGCCAC 387
QY 121 AGATCTAATGCTCTTCACTGAGTGGGCGGAAACATTAATTAATGAGCAAGCAAC 180
DB 388 AGATCTAATGCTCTTCACTGAGTGGGCGGAAACATTAATTAATGAGCAAGCAAC 447
QY 181 CGAGGCGGAGCCGCTGCTATGTCAGTGGGCGCTTAAAGCCGCTTCCAGAGTGCATG 240
DB 448 CGAGGCGGAGCCGCTGCTATGTCAGTGGGCGCTTAAAGCCGCTTCCAGAGTGCATG 507
QY 241 GTGCATGACGCGCAGAT 258
DB 508 GTGCATGACGCGCAGAT 525
```

```

RESULT 14
US-09-954-456-552
; Sequence 552, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
```

```

; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent version 3.0
; SEQ ID NO: 552
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-552

Query Match
Best Local Similarity 27.4%; Score 70.8; DB 10; Length 2036;
Matches 138; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```

QY 6 CTGCTATGAGGGGATGGTCACTTTACCGAGAAAGCCGACACTGACACCATGAGGCGG 65
DB 858 CTGCTATGAGGGGATGGTCACTTTACCGAGAAAGCCGACACTGACACCATGAGGCGG 917
QY 66 GCGCTGCTGCTGCTGGAACCTGCGCACTGCTTCCAGCAACGTAACATGCCACAGATC 125
DB 918 CAGCTGCTGCTGCTGGAACCTGCGCACTGCTTCCAGCAACGTAACATGCCACAGATC 977
QY 126 TAATGCTCTTCACTGAGTGGGCGGAAACATTAATTAATGAGCAAGCAACCGGAG 185
DB 978 GCGCTGCTGCTGCTGGAACCTGCGCACTGCTTCCAGCAACGTAACATGCCACAGATC 1037
QY 186 GCGACCCGAGTGTATGTCAGTGGGCGCTTAAAGCCGCTTCCAGAGTGCATGATGCA 245
DB 1038 GAGGCGCTGCTGCTGTCAGTGGGCGCTTAAAGCCGCTTCCAGAGTGCATGATGCA 1097
QY 246 TGACTGCGCA 255
DB 1098 GGCTGCGCA 1107
```

```

RESULT 15
US-09-880-107-1612
; Sequence 1612, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US/60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US/60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO: 1612
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D14012
US-09-880-107-1612
```

```

Query Match
Best Local Similarity 27.4%; Score 70.8; DB 10; Length 2036;
Matches 138; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 6 CTGCTATGAGGGGATGGTCACTTTACCGAGAAAGCCGACACTGACACCATGAGGCGG 65
DB 858 CTGCTATGAGGGGATGGTCACTTTACCGAGAAAGCCGACACTGACACCATGAGGCGG 917
QY 66 GCGCTGCTGCTGCTGGAACCTGCGCACTGCTTCCAGCAACGTAACATGCCACAGATC 125
```





1  
2  
3  
4  
5



/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="387620"  
 /clone\_lib="NIH\_MGC\_69"  
 /issue\_type="large cell carcinoma, undifferentiated"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: Lung; Vector: PCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.1 kb. Library constructed by Life  
 Technologies."  
 BASE COUNT 188 a 185 c 184 g 143 t  
 ORIGIN

Query Match 99.4%; Score 256.4; DB 12; Length 700;  
 Best Local Similarity 99.6%; Pred. No. 4.8e-67;  
 Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAACCTGATGAGGGGATGTCACCTTTACCGAGAAAGCCAGCAGTACACCATG 60  
 |||||||  
 Db 118 AAAACCTGATGAGGGGATGTCACCTTTACCGAGAAAGCCAGCAGTACACCATG 177  
 |||||||  
 OY 61 GCGCGCCCTGCGCTGCGCACTGCTCTTACGAAACGATGATGCTCCAC 120  
 |||||||  
 Db 178 GCGCGCCCTGCGCTGCGCACTGCTCTTACGAAACGATGATGCTCCAC 237  
 |||||||  
 OY 121 AGATCTAATGCTCTTACGCTGCGCTGCGCACTGCTCTTACGAAACGATGATGCTCCAC 180  
 |||||||  
 Db 238 AGATCTAATGCTCTTACGCTGCGCTGCGCACTGCTCTTACGAAACGATGATGCTCCAC 297  
 |||||||  
 OY 181 CGAGGCGACCCCTGCTGATGTCAGAGTGGGCTTAAAGCCGCTTGCAGAGTGCATG 240  
 |||||||  
 Db 298 CGAGGCGACCCCTGCTGATGTCAGAGTGGGCTTAAAGCCGCTTGCAGAGTGCATG 357  
 |||||||  
 OY 241 GTGCATGACTGCGCAGAT 258  
 |||||||  
 Db 358 GTGCATGACTGCGCAGAT 375  
 |||||||

RESULT 2  
 BE877571 767 bp mRNA linear EST 20-OCT-2000  
 LOCUS BE877571  
 DEFINITION 601486523p1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3888964 5',  
 mRNA sequence.  
 ACCESSION BE877571  
 VERSION BE877571.1 GI:10326347  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 767)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rcmail.nih.gov  
 Tissue Procurement: DCTD/DP/Genetic  
 CNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LLM6669 row: n column: 05  
 High quality sequence stop: 710.  
 Location/Qualifiers

FEATURES  
 source  
 1..767  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="3888964"  
 /clone\_lib="NIH\_MGC\_69"  
 /issue\_type="large cell carcinoma, undifferentiated"  
 /lab\_host="DH10B (phage-resistant)"

/note="Organ: Lung; Vector: PCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.1 kb. Library constructed by Life  
 Technologies."  
 BASE COUNT 183 a 228 c 204 g 152 t  
 ORIGIN

Query Match 99.4%; Score 256.4; DB 12; Length 767;  
 Best Local Similarity 99.6%; Pred. No. 5.1e-67;  
 Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAACCTGATGAGGGGATGTCACCTTTACCGAGAAAGCCAGCAGTACACCATG 60  
 |||||||  
 Db 288 AAAACCTGATGAGGGGATGTCACCTTTACCGAGAAAGCCAGCAGTACACCATG 347  
 |||||||  
 OY 61 GCGCGCCCTGCGCTGCGCACTGCTCTTACGAAACGATGATGCTCCAC 120  
 |||||||  
 Db 348 GCGCGCCCTGCGCTGCGCACTGCTCTTACGAAACGATGATGCTCCAC 407  
 |||||||  
 OY 121 AGATCTAATGCTCTTACGCTGCGCTGCGCACTGCTCTTACGAAACGATGATGCTCCAC 180  
 |||||||  
 Db 408 AGATCTAATGCTCTTACGCTGCGCTGCGCACTGCTCTTACGAAACGATGATGCTCCAC 467  
 |||||||  
 OY 181 CGAGGCGACCCCTGCTGATGTCAGAGTGGGCTTAAAGCCGCTTGCAGAGTGCATG 240  
 |||||||  
 Db 468 CGAGGCGACCCCTGCTGATGTCAGAGTGGGCTTAAAGCCGCTTGCAGAGTGCATG 527  
 |||||||  
 OY 241 GTGCATGACTGCGCAGAT 258  
 |||||||  
 Db 528 GTGCATGACTGCGCAGAT 545  
 |||||||

RESULT 3  
 AUI31687 799 bp mRNA linear EST 01-AUG-2002  
 LOCUS AUI31687  
 DEFINITION NT2RP3 Homo sapiens cDNA clone NT2RP3003050 5', mRNA  
 sequence.  
 ACCESSION AUI31687  
 VERSION AUI31687.1 GI:10992041  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 799)  
 Ota, T., Nishikawa, Y., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,  
 Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and  
 Isogai, T.  
 HRI human cDNA project  
 Unpublished (2000)  
 Contact: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project: 5'-, 3'-end one pass sequencing: Helix  
 Research Institute; cDNA library construction: Department of  
 Virology, Institute of Medical Science, University of Tokyo, and  
 Helix Research Institute  
 Location/Qualifiers

FEATURES  
 source  
 1..799  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="NT2RP3003050"  
 /clone\_lib="NT2RP3"  
 /cell\_type="teratocarcinoma"  
 /cell\_line="NT2"  
 /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor  
 cells after 2-weeks retinoic acid (RA) induction"  
 BASE COUNT 198 a 233 c 212 g 152 t 4 others  
 ORIGIN



Tue Nov 12 18:06:53 2002

pct-us02-27855-2.rst

Page 4

	Sequence	Position
OY	1 AAAACCTGCTATGAGGGGAATGTCACCTTTACCGGAGAAAGCCGACACTACACATG	60
Db	308 AAAACCTGCTATGAGGGGAATGTCACCTTTACCGGAGAAAGCCGACACTACACACATG	367
OY	61 GGGCGGCGCTGGCTGCCCGGAACCTGCGACTGTCCTTCAGCAAAAGTAACATGCGCAC	120
Db	368 GGGCGGCGCTGGCTGCCCGGAACCTGCGACTGTCCTTCAGCAAAAGTAACATGCGCAC	427
OY	121 AGATCTAATGCTCTTCACGCTGGGCGCTGGGGAAACATAATTACTGCAGAGAACCCAGCAAC	180
Db	428 AGATCTAATGCTCTTCACGCTGGGCGCTGGGGAAACATAATTACTGCAGAGAACCCAGCAAC	487
OY	181 CGGAGGCGCACTGTGTCTATGTGCAAGGTGGGCGCTTAAAGCCGCTTGCCAAAGAGTGCAATG	240
Db	488 CGGAGGCGCACTGTGTCTATGTGCAAGGTGGGCGCTTAAAGCCGCTTGCCAAAGAGTGCAATG	547
OY	241 GTGCATGACTGCGCAGAT	258
Db	548 GTGCATGACTGCGCAGAT	565

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 913)	NIH-MGC	<a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a>	National Institutes of Health, Mammalian Gene Collection (MGC)	unpublished (1999)
	Contact: Robert Strausberg, Ph.D.			

Email: [cgapbbs-remail.nih.gov](mailto:cgapbbs-remail.nih.gov)  
Tissue Procurement: DCTD/DFP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLNL at:  
<http://image.llnl.gov>  
Plate: LICM2568 row: a column: 14  
High quality sequence stop: 595.  
Location/Qualifiers

FEATURES	source
Location/Qualifiers	1. 913
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:6380053"	
/clone_lib="NH_MGC_40"	
/tissue_type="carcinoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: prostate; Vector: pOT8; Site:1: XhoI site; 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis (Stratagene) and Superscript II RT (Life Technologies)." Note: this is a NIH_MGC Library.	
BASE COUNT	217 a 276 c 250 g 169 t 1 others
ORIGIN	

Query Match	99.4%	Score 256.4	DB 14	Length 913
Best Local Similarity	99.6%	Pred No.5-5e-67		
Matches 257	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	AAAACCGCCTATAGGGGAATGTCACCTTTTACCGAGAAAGGCCACACTGCACCATG	60	

Db	310	AAAACTCTTAAGAGGGGAATGTCACTTTTACCGAGGAAAGGCGACACTGATGACCATG	366
Qy	61	GGCGCGCCCTGCTGCTCCCTGGAACTCTCCACTGTCCTTCAGCAAGAGTACATGCCCAC	126
Db	370	GGCGGCGCCCTGCTGCTCCCTGGAACTCTGCCACTGTCTTCAGCAAGAGTACATGCCCAC	428
Qy	121	AGATCTAATGCTCTTCACACTGGGCTGGGGAAACATAATTACTGACAGAACCCAGACAC	180
Db	430	AGATCTGATGCTCTTCACACTGGGCTGGGGAAACATAATTACTGACAGAACCCAGACAC	489
Qy	181	CGGAGCGCAACCTGGTGTATATGTCAGATGGGCTTAAAGCCGCTTGTCGAAGATGTCATG	240
Db	490	CGGAGCGCAACCTGGTGTATATGTCAGATGGGCTTAAAGCCGCTTGTCGAAGATGTCATG	549
Qy	241	GTGCATGACTGGGAGAT	258
Db	550	GTGCATGACTGGGAGAT	567

RESULT 7	930 bp	mRNA	linear	EST_15-SEP-2000
BE742186				
LOCUS				
DEFINITION	BE742186	6015757584p1	NIH_MGC_9	Homo sapiens CDNA clone IMAGE:583612 5',
				mRNA sequence.
ACCESSION	BE742186			
VERSION	BE742186.1	GI:10156178		
KEYWORDS	EST.			
SOURCE	human			
ORGANISM				

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 930)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: cgabps-r@mail.nih.gov  
Tissue Procurement: DCTD/DRP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
Library Screening: J. M. A. G. E. Consortium (LNLN)  
DNA Sequencing: J. M. A. G. E. Consortium  
Clone distribution: MGC clone distribution information can be found through the I. M. A. G. E. Consortium/LNLN at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LHC520 row: h column: 21  
High quality sequence stop: 743.

FEATURES	Location/Qualifiers
SOURCE	1. 930
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone_image:3636612"
	/clone_lib="NH_MGC_9"
	/tissue_type="adenoecarcinoma cell line"
	/lab_host="DH10B (phage-resistant)"
	/note="Organ: ovary; Vector: pORF7. Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT	217 a 265 c 264 g 184 t
ORIGIN	

```

Query Match      99.4%   Score 256.4; DB 12; Length 930;
Best Local Similarity 99.6%; Pred No. 5,6e-67;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY    1 AAAAAGCTGCAATGACGGGAATGTACTTTTACGAGAAGGCCACACTGACCACATG 60
      |||||
Db     20 AAAAACCTGCATATGAGGGGAAATGTACTTTTACCGAGAAGGCCACACTGACCCATTG 79

```

QY	61	GGCGGGCCCTCTCGTGGCCCTGGAACTGTGCACATGCTCCTTAGCAAAAGTCACATGGCCAC	120
Db	80	GGCGGGCCCTCTCGTGGCCCTGGAACTGTGCACATGCTCCTTAGCAAAAGTCACATGGCCAC	139
QY	121	AGATCTAATGCTCTTTCAGCTGGGCGCTGGGAAACATATTAATCTACAGAACCCAGAAC	180
Db	140	AGATCTGATGCTTTCAGAGCTGGGCGCTGGGAAACATATTAATCTACAGAACCCAGAAC	199
QY	181	CGGAGGCGACCTCTGTGCTATGTGCAGGTGGGCGCTTAAGCGGCTTGTCGAAGATGCATG	240
Db	200	CGGAGGCGACCTCTGTGCTATGTGCAGGTGGGCGCTTAAGCGGCTTGTCGAAGATGCATG	259
QY	241	GTGCATGACCTCGCAGAT	258
Db	260	GTGCATGACCTCGCAGAT	277

RESULT 8	
BQ961887	
LOCUS	932 bp mRNA linear
DEFINITION	AGNCCOURT_8778892 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6378687 5' , mRNA sequence.

VERSION	BO961887.1	GI:22377365
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
REFERENCE	1 (bases 1 to 932)	
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished. (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D.	

Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 plate: LHC8564 row: h column: 16  
 High quality sequence stop: 606.  
 Location/Qualifiers  
 I. .932

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6378687"
/clone_1fb="NIH_MGC_42"
/tissue_type="epithelial carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pOTB3; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library. |"

```

	Query Match	Best Local Similarity	Matches	Conservative	Score	DB	Length	932:
0y	1	AAAACCTGCTATGAGGGGATGTCATCTTTATCCGAGAAAGGCACGACGACACCTG	60	0	Indels	0	Gaps	0;
Db	303	AAAACCTGCTATGAGGGGATGTCATCTTTATCCGAGAAAGGCACGACGACACCTG	362					
0y	61	GCGCGGCGCTGCTGCGCTCGACACTTCCCACTGTCTTTCAGCAAAACGTACCATGCCCAC	120					

Accession	Sequence	Length
Db	363 GGCGGCCCTCCTCCCTGGAACTCTGCACCTCCCTTCAGCAAAAGTACCAAGCCAC	422
Qy	121 AGATTAATAGCTCTACACTGGGCGCTGGGAAACATAATTACTGACAGAACCCAGACAC	180
Db	423 AGATCTGATGCTCTTCACACTGGGCGCTGGGAAACATAATTACTGACAGAACCCAGACAC	482
Qy	181 CGSAGCGCACCCCTGGTGTATGTGCAGGTGGGCGCTAAAGCGGCTGTCTCAAGAGTGCATG	240
Db	483 CGSAGCGCACCCCTGGTGTATGTGCAGGTGGGCGCTAAAGCGGCTGTCTCAAGAGTGCATG	542
Qy	241 GTGCATGACTCGCGCAGAT	258
Db	543 GTGCATGACTCGCGCAGAT	560

RESULT 9					
B0929215					
LOCUS	B0929215	963 bp	mRNA	linear	EST 20-AUG-2002
DEFINITION	AGENCOURT 8958654 NIH_MGC_101 Homo sapiens CDNA clone IMAGE:6460542				
ACCESSION	5', mRNA sequence.				
	3'				

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 963)
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

```

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1CM2647 row: g column: 07
High quality sequence stop: 566.
Location/Qualifiers
1. .963
FEATURES
Source

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6460542"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Lung; Vector: pOTB7; Site_1: EcoRI, Site_2:
XhoI; DNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: This is a
NIH_MGC Library."

```

	Query Match	99.48%	Score 256.4	DB 14	Length 963
	Best Local Similarity	99.6%	Pred. No. 5.7e-67		
	Matches 257	Conservative	0	Mismatches 1	Indels 0
				Gaps	0
QY	1	AAACCCGCTATAGGAGGAATGGTCACTTTACGACGAGAAGGCCACACTGACACCATG	60		
DB	277	AAACCTCGCATATAGGAGGATGGTCTCTTTACGAGGAAAAGGCCACACTGACACCATG	336		
QY	61	GAGCGGCGCTGTCGTCGCGCTGGAACCTGTGACACTGTCTTTCAGCAAAAGTCACATGGCCAC	120		
DB	337	GAGCGGCGCTGTCGTCGCGCTGGAACCTGTGACACTGTCTTTCAGCAAAAGTCACATGGCCAC	396		

Tue Nov 12 18:06:53 2002

pct-us02-27855-2.rst

Page 6

Qy	121	AGATTTATTTGCTTCCTCAGCTGGGACCTGGGGAAACATATTTTCTGCGAGAACCCAGACAC	180
Db	397	AGATCTGATGCTCTTCACACTGGGCTGGGGAAACATATTTACTGCGAGAACCCAGACAC	456
Qy	181	CGGAGGCGAACCCCTGGTGGTATGTGAGAGTGGGCGCTTAAAGCCGCTGTCACAGAGTCATG	240
Db	457	CGGAGGCGAACCCCTGGTGGTATGTGAGAGTGGGCGCTTAAAGCCGCTTCCAGAGTCATG	516
Qy	241	GTGCATGACTGGCGCAGAT	258
Db	517	GTGCATGACTGGCGCAGAT	534

RESULT	10
LOCUS	BQ922207
DEFINITION	BQ922207 1001 bp mRNA linear EST 20-AUG-2002
ACCESSION	AGNCCOURT.8945721 NIH_MGC_101 Homo sapiens CDNA clone IMAGE:6459780
KEYWORDS	5' mRNA sequence.
SOURCE	BQ922207 BQ922207.1 GI:22337238
ORGANISM	EST. human. Homo sapiens Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1001)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: [cgapso-remail@nlb.gov](mailto:cgapso-remail@nlb.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
 row: 9 column: 13  
 Plate: LLCM2645 spot: 520.  
 High quality sequence

## FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6459780"
/clone_lib="NIH.MGC.101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Lung; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH.MGC Library."

```

	Query Match	Similarity	Best Local	Score	DB	Length
Matches	257	Conservative	0	Mismatches	1	Gaps
QY	1	AAAACCTGCTATGAGGGGGAATGCTACATTTTTCGAGGAAAAGCCAGCACTACACCATG	60			
DB	277	AAAACCTGCTATGAGGGGGAATGCTACATTTTTCGAGGAAAAGCCAGCACTACACCATG	336			
QY	61	GGCGGCGCCCTGCTGCCCTCGAAGCACTGCACATGTCCTTCAGCAAAAGACATGCCAC	120			
DB	337	GGCGGCGCCCTGCTGCCCTCGAAGCACTGCACATGTCCTTCAGCAAAAGACATGCCAC	396			
QY	121	AGATCTAATGCTCTTTCAGCTGGGCGCTGGGGAACAATAATTACTGACGAGAACCCAGACAA	180			

Accession	Sequence	Length
Db	AGATCTGATGCTCTTCAAGCTGGGCGCTGGGGAAACATATTAATGACAGAACCCAGAAC	456
Oy	CGAAGCGCACCCCTGGTGTATGTATGCAGGTGGGCGCTTAAAGCGCGTTGTCGAAGATGCATG	240
Db	CGAAGCGCACCCCTGGTGTATGTATGCAGGTGGGCGCTTAAAGCGCGTTGTCGAAGATGCATG	516
Oy	GTGCATGACTCGCGAGAT	258
Db	GTGCATGACTCGCGAGAT	534

RESULT	11			
Bm903637				
LOCUS		1063 bp	mRNA	linear EST 12-MAR-2002
DEFINITION	Bm903637			
	AGGCCOURT_6622158 NIH_MGC_125 Homo sapiens CDNA clone IMAGE:5722049			
	5' mRNA sequence.			
ACCESSION	Bm903637			
VERSION	Bm903637.1	GI:19353039		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

REFERENCE	Eumaiyotia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
AUTHORS	1 (bases 1 to 1065)
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: [cgadps+email@nlb.gov](mailto:cgadps+email@nlb.gov)  
Tissue Procurement: Invitrogen  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LVM42707 row: h column: 18  
High quality sequence spot: 742.

**FEATURES**  
**Source**

BASE COUNT 257 a 315 c 282 g 208 t 1 others  
ORIGIN

Query Match	Best Local Similarity	Score	DB	Length
Matches	257; Conservative	99.4%; 99.6%;	0; Mismatches	0; Gaps
0Y	1	AAAACCTGTTAGAGGGGAAATGGTACATTTTACCGAGGAAAGGACGACACTGACACATG	60	
Db	311	AAAACCTGTTAGAGGGGAAATGGTACATTTTACCGAGGAAAGGACGACACTGACACATG	370	
QY	61	GGCGGGCCCTGGCTGCGCCCGGAACCTGCGACATGTGCTTCGAGCAACGTACCATGCGAC	120	
Db	371	GGCGGGCCCTGGCTGCGCCCGGAACCTGCGACATGTGCTTCGAGCAACGTACCATGCGAC	430	
QY	121	AGATCTAATGCTCTTTCACGCTGGGGCTGGGGAAACATAATTTACTGAGGAAACCCAGACAC	180	
Db	431	AGATCTAATGCTCTTTCACGCTGGGGCTGGGGAAACATAATTTACTGAGGAAACCCAGACAC	490	



Accession	Sequence	Position
Qy	CGAGGCGACACCTGGTGGTATGTACAGTGGGCTAAAGCCCTGTCCAGAGATGATG	240
Db	CGAGGCGACACCTGGTGGTATGTACAGTGGGCTAAAGCCCTGTCCAGAGATGATG	550
Qy	GTGATGACTGCGGCAGAT	258
Db	GTGATGACTGCGGCAGAT	568

RESULT 12					
BM920114					
LOCUS	1067 bp	mRNA	linear	EST 12-MAR-2007	
DEFINITION	BM920114				
	AGENCOURT-6706522	NIH_MGC_120	Homo sapiens	cDNA clone IMAGE:574983	
	5', mRNA sequence.				

ACCESSION	BM920114
VERSION	BM920114.1
KEYWORDS	EST.
SOURCE	human.

ORGANISM	REFERENCE
<i>Homo sapiens</i>	1 (bases 1 to 1067)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
RTM-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a>	
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	
Contact: Robert Strausberg, Ph.D.	

Email: [cgabbs-remail@lin.gov](mailto:cgabbs-remail@lin.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM12780 row: d column: 20  
 High quality sequence stop: 754.

## FEATURES

## Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5749987"
/clone_1ib="NIH_MGC_120"
/lab_host="DH10B"
/notes="Organ: pooled pancreas and spleen; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
pool= anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb. Insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

```

Query Match	99.4%;	Score 256.4;	DB 14;	Length 1067;
Best Local Similarity	99.6%;	Pred. No. 6e-67;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

Oy	1	AAAACCGCTATGAGGGGAATGTCATCTTTACCGAGGAAAGCCAGCACTGACACCATG	60
Db	328	AAAACCTGCTATGAGGGGAATGTCATCTTTACCGAGGAAAGCCAGCACTGACACCATG	387
Oy	61	GGCGCGCCCTGCTGCGCTCGAAGCTCTGCCACTGTCCCTTAGCAAAAGTACCATGCCCAC	120
Db	388	GGCGCGCCCTGCTGCGCTCGAAGCTCTGCCACTGTCCCTTAGCAAAAGTACCATGCCCAC	447
Oy	121	AGATCTATAGCTCTTCAGCTGGGGCTGGGGAAACATATTTACTGCAGGAACCCAGAACAC	180
Db	448	AGATCTATGCTCTTCAGCTGGGGCTGGGGAAACATATTTACTGCAGGAACCCAGAACAC	507
Oy	181	CGGAGCGCAACCTTGCTCTATGTGCAAGGTGGGGCTTAAGCCGCTTGTCCAAAGTGCATG	240

Db 508 CCGAAGCGCAACCCCTGGTGCCTATGTGTCAGGTGGCCCTAAAGCCGCTGTCCAGAGATGCATG 567

Qy 241 GTGCATGACTGCGCAGAT 258

Db 568 GTGCATGACTGCGCAGAT 585

RESULT 13	BE742275	1075 bp	mRNA	linear	EST 15-SEP-2000
LOCUS	601575563F1				
DEFINITION	Homo sapiens CDNA clone IMAGE:3836663 5', mRNA sequence.				

ACCESSION	BE742275	GI:10156267
VERSION	BE742275.1	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.  
1 (bases 1 to 1075)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: [cgap@rs-internal.nih.gov](mailto:cgap@rs-internal.nih.gov)  
 Tissue Procurement: DCTD/DRP  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: LNCM520 row: j column: 24  
 High quality sequence stop: 681

## FEATURES

### SOURCE

## Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3836663"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/not=Organ: ovary; Vector: pORF7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

QY 1 AAAACCTGATGAGGGGAATGTCACTTTTACCAGAGAAGGCCAGCACTGCACCATTG 60  
Db 20 AAACCTGATGAGGGGAATGTCACTTTTACCAGAGAAGGCCAGCACTGCACCATTG 79

QY	61	GGCCGGCCCTGCGCTGCGCAACTCTGCCACTGCTCTTCAGCAACAGTACCATGCCCC	120
	80	GGCCGGCCCTGCGCTGCGCAACTCTGCACTGCTCTTCAGCAACAGTACCATGCCCC	139
Db			
QY	121	AGACTATATCTCTTTAGCTGGGCGCTGGGGAAACATATATTACGACGAGAACCCAGACAC	180
	140	AGACTATATCTCTTTAGCTGGGCGCTGGGGAAACATATATTACGACGAGAACCCAGACAC	199
Db			
QY	181	CGAGGGGACCGCTGGTGTATGTGCAAGTGGGCGCTTAAGCGCGTTTCCAGAGTGCATG	240
	200	CGAGGGGACCGCTGGTGTATGTGCAAGTGGGCGCTTAAGCGCGTTTCCAGAGTGCATG	255
Db			

Tue Nov 12 18:06:53 2002

pct-us02-27855-2.1st

Page 8

```

QY 241 GTGCATGACTGGCAGAT 258
|||||
Db 260 GTGCATGACTGGCAGAT 277

RESULT 14
LOCUS BE897378
DEFINITION 601437369F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3922288 5',
mRNA sequence.
ACCESSION BE897378
VERSION BE897378.1 GI:10362780
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 1087)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCMD/DPF
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM9756 row: 1 column: 17
High quality sequence stop: 654.
Location/Qualifiers
1. 1087
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3922288"
/clone_lib="NIH-MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1: Not;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo-dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 273 a 345 c 268 g 200 t 1 others
ORIGIN
Query Match 99.4%; Score 256.4; DB 12; Length 1087;
Best Local Similarity 99.6%; Pred. No. 6e-67; Indels 0; Gaps 0;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGATGCTACTTTTACCGAGAAAGGCCAGCTGACACCATG 60
|||||
Db 268 AAAACCTGCTATGAGGGGATGCTACTTTTACCGAGAAAGGCCAGCTGACACCATG 327

QY 61 GGGCGGCGCTGCTGCGCTGGAAGTCTGCACTGCTTACGAGAAAGGTCATGCGCCAC 120
|||||
Db 328 GGGCGGCGCTGCTGCGCTGGAAGTCTGCACTGCTTACGAGAAAGGTCATGCGCCAC 387

QY 121 AGATCTAATGCTCTTCACTGAGCTGGGCGGGAACATTAATTAAGCAAGCCAGAAC 180
|||||
Db 388 AGATCTAATGCTCTTCACTGAGCTGGGCGGGAACATTAATTAAGCAAGCCAGAAC 447

QY 181 CGAGAGGCGACCTGCTGCTATGTCAGAGTGGGCTTAAAGCCGCTTGTCCAAAGTGCATG 240
|||||
Db 448 CGAGAGGCGACCTGCTGCTATGTCAGAGTGGGCTTAAAGCCGCTTGTCCAAAGTGCATG 507

QY 241 GTGCATGACTGGCAGAT 258
|||||
Db 508 GTGCATGACTGGCAGAT 525
```

```

RESULT 15
LOCUS BF568088
DEFINITION 602184082F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:429981 5',
mRNA sequence.
ACCESSION BF568088
VERSION BF568088.1 GI:11641506
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 1164)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM1158 row: c column: 22
High quality sequence stop: 727.
Location/Qualifiers
1. 1164
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="429981"
/clone_lib="NIH-MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pORF7; Site: 1: XhoI;
Site: 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(C). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."
BASE COUNT 264 a 361 c 324 g 215 t
ORIGIN
Query Match 99.4%; Score 256.4; DB 12; Length 1164;
Best Local Similarity 99.6%; Pred. No. 6.3e-67; Indels 0; Gaps 0;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGATGCTACTTTTACCGAGAAAGGCCAGCTGACACCATG 60
|||||
Db 308 AAAACCTGCTATGAGGGGATGCTACTTTTACCGAGAAAGGCCAGCTGACACCATG 367

QY 61 GGGCGGCGCTGCTGCGCTGGAAGTCTGCACTGCTTACGAGAAAGGTCATGCGCCAC 120
|||||
Db 368 GGGCGGCGCTGCTGCGCTGGAAGTCTGCACTGCTTACGAGAAAGGTCATGCGCCAC 427

QY 121 AGATCTAATGCTCTTCACTGAGCTGGGCGGGAACATTAATTAAGCAAGCCAGAAC 180
|||||
Db 428 AGATCTAATGCTCTTCACTGAGCTGGGCGGGAACATTAATTAAGCAAGCCAGAAC 487

QY 181 CGAGAGGCGACCTGCTGCTATGTCAGAGTGGGCTTAAAGCCGCTTGTCCAAAGTGCATG 240
|||||
Db 488 CGAGAGGCGACCTGCTGCTATGTCAGAGTGGGCTTAAAGCCGCTTGTCCAAAGTGCATG 547

QY 241 GTGCATGACTGGCAGAT 258
|||||
Db 548 GTGCATGACTGGCAGAT 565
```

Search completed: November 11, 2002, 12:57:41